

B90

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A2	(11) International Publication Number: <b>WO 99/65938</b>
C07K 14/00		(43) International Publication Date: 23 December 1999 (23.12.99)

(21) International Application Number: PCT/US99/13741	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 17 June 1999 (17.06.99)	
(30) Priority Data: 60/089,556 17 June 1998 (17.06.98) US	
(71) Applicant: AMERSHAM PHARMACIA BIOTECH INC. [US/US]; 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US).	
(72) Inventors: DAVIS, Maria; Amersham Pharmacia Biotech Inc., 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US). FULLER, Carl, W.; Amersham Pharmacia Biotech Inc., 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US). MAMONE, Joseph, A.; Amersham Pharmacia Biotech Inc., 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US). HUANG, Lin; Amersham Pharmacia Biotech Inc., 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US).	
(74) Agent: RONNING, Royal, N., Jr.; Amersham Pharmacia Biotech Inc., 800 Centennial Avenue, P.O. Box 1327, Piscataway, NJ 08855-1327 (US).	

(54) Title: FY7 POLYMERASE

## (57) Abstract

A purified recombinant thermostable DNA polymerase which exhibits at least about 80 % activity at salt concentrations of 50 mM and greater, at least about 70 % activity at salt concentrations of 25 mM and greater, and having a processivity of about 30 nucleotides per binding event. An isolated nucleic acid that encodes the thermostable DNA polymerase, as well as a recombinant DNA vector comprising the nucleic acid and a recombinant host cell transformed with the vector, are also disclosed. A method of sequencing DNA using the DNA polymerase as well as a kit for sequencing DNA is also disclosed.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## FY7 POLYMERASE

### CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to United States Provisional Application Serial No. 60/089,556, filed on June 17, 1998, the entire disclosure of which is incorporated in its herein.

### BACKGROUND OF THE INVENTION

#### Field of the Invention

The instant disclosure pertains to thermostable DNA polymerases which exhibit improved robustness and efficiency.

#### Background

DNA polymerases are enzymes which are useful in many recombinant DNA techniques such as nucleic acid amplification by the polymerase chain reaction ("PCR"), self-sustained sequence replication ("3SR"), and high temperature DNA sequencing. Thermostable polymerases are particularly useful. Because heat does not destroy the polymerase activity, there is no need to add additional polymerase after every denaturation step.

However, many thermostable polymerases have been found to display a 5' to 3' exonuclease or structure-dependent single-stranded endonuclease ("SDSSE") activity which may limit the amount of product produced or contribute to the plateau phenomenon in the normally exponential accumulation of product. Such 5' to 3' nuclease activity may contribute to an impaired ability to efficiently generate long PCR products greater than or equal to 10kb, particularly for G+C rich targets. In DNA sequencing applications and cycle sequencing applications, the presence of 5' to 3' nuclease activity may contribute to a reduction in desired band intensities and/or generation of spurious or background bands.

Additionally, many of the enzymes presently available are sensitive to high salt environments, a condition commonly

Presently available enzymes have so-so processing ability (are more distributive - fall off more often – explain in more detail)

dITP added to address compression problems – usually kills activity of enzyme

Thus, a need continues to exist for an improved DNA polymerase having increased tolerance to high salt conditions, efficient utilization of dITP, high productivity, and improved performance on GC-rich templates.

### BRIEF SUMMARY OF THE INVENTION

The instant disclosure teaches a purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 1, as well as a purified recombinant thermostable DNA polymerase which exhibits at least about 80% activity at salt concentrations of 50 mM and greater. The instant disclosure further teaches a purified recombinant thermostable DNA polymerase which exhibits at least about 70% activity at salt concentrations of 25 mM and greater, and a purified recombinant thermostable DNA polymerase having a processivity of about 30 nucleotides per binding event.

The instant disclosure also teaches an isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence set forth in Figure 1, as well as a recombinant DNA vector that comprises the nucleic acid, and a recombinant host cell transformed with the vector.

The instant disclosure also teaches a method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments. The instant disclosure also teaches a kit for sequencing DNA comprising the DNA polymerase.

### BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

FIGURE 1 depicts the amino acid sequence (and DNA sequence encoding therefor) for the FY7 polymerase.

FIGURE 2 depicts the DNA sequence of M13mp18 DNA sequenced using the FY7 polymerase formulated in Mn conditions, as shown by a print out from an ABI model 377 automated fluorescent DNA sequencing apparatus.

FIGURE 3 depicts the DNA sequence of M13mp18 DNA sequenced using the FY7 polymerase formulated in Mg conditions, as shown by a print out from an ABI model 377 automated fluorescent DNA sequencing apparatus.

FIGURE 4 depicts the percent of maximum polymerase activity for Thermo Sequenase<sup>TM</sup> enzyme DNA polymerase versus FY7 DNA polymerase under varying KCl concentrations.

FIGURE 5 depicts the effect of high salt concentrations on DNA sequencing ability in radioactively labeled DNA sequencing reactions using Thermo Sequenase™ enzyme DNA polymerase versus FY7 DNA polymerase.

FIGURES 6-10 depict the effect of increasing salt concentration on the performance of Thermo Sequenase. At concentrations as low as 25mM data quality is affected with the read length being decreased from at least 600 bases to about 450 bases. At 50mM salt the read length is further decreased to about 350 bases, 75mM to about 250 bases and at 100mM the read length is negligible.

FIGURES 11-15 depict the effect of increasing salt concentration on the performance of FY7 DNA polymerase. There is no detrimental effect on performance to at least 75mM KCl and only a slight decrease in data quality at 100mM KCl.

FIGURE 16 depicts the processivity measured for Thermo Sequenase DNA polymerase, AmpliTaq FS DNA polymerase, compared with the processivity measured for FY7 DNA polymerase.

FIGURE 17 depicts the improved read length obtained when using FY7 polymerase versus Thermo Sequenase DNA polymerase in radioactively labeled sequencing reactions incorporating the dGTP (Guanosine triphosphate) analog dITP (Inosine triphosphate) at 72 °C.

FIGURES 18-22 show the effect of increasing extension step time on the read length and data quality produced by Thermo Sequenase DNA polymerase in fluorescently labeled terminator DNA sequencing reactions

FIGURES 23-27 show the effect of increasing extension step time on the read length and data quality produced by FY7 DNA polymerase in fluorescently labeled terminator DNA sequencing reactions.

#### DETAILED DESCRIPTION OF THE INVENTION

A series of polymerase mutants were constructed with the aim of obtaining an improved polymerase for DNA sequencing, by reducing the exonuclease activity found in full length *Thermus thermophilus* and *Thermus aquaticus* DNA polymerase I enzymes. Six conserved motifs (Gutman and Minton (1993) Nucleic Acids Research 21, 4406 - 4407) can be identified in the amino-terminal domain of pol I type polymerases, in which the 5' to 3' exonuclease activity has been shown to reside. Further, six carboxylate residues in these conserved regions have been shown in a crystal structure to be located at the active site of the exonuclease domain of *Thermus aquaticus* DNA pol I (Kim et al., (1995) Nature 376, 612-616). Point mutations were made by site-directed mutagenesis to carboxylates and other residues in three of six conserved motifs in Tth and Taq polymerases as follows:

Taq D18A, Taq T140V, Taq D142N/D144N. All of these have the mutation F667Y outside of the exonuclease domain.

Tth D18A, Tth T141V, Tth D143N/D145N. All of these have the mutation F669Y outside of the exonuclease domain.

All polymerases were evaluated for exonuclease activity, processivity, strand displacement, salt tolerance, thermostability, and sequencing quality. One FY7 polymerase, Tth D18A, F669Y, is described in further detail below.

## EXAMPLES

### Methods

#### In vitro mutagenesis

PCR was employed to introduce an aspartic acid to alanine amino acid change at codon 18 (D18A) of cloned full length F669Y Tth (plasmid pMR10). Mutagenic Primer 1 (CTGTTCGAACCAAAGGCCGTGTCCTCCTGGTGGCCGGCCACCA) spans nucleotides 19-60 of pMR10 including codon 18 and a *Bst*BI restriction site. Oligonucleotide Primer 2 (GAGGCTGCCGAATTCCAGCCTCTC) spans an *Eco*RI site of pMR10. pMR10 was used as template DNA. The PCR product was digested with *Bst*BI and *Eco*RI and ligated to two fragments of pMR10: a 5000 bp *Kpn*I/ *Bst*BI and a 2057 bp *Eco*RI / *Kpn*I, creating plasmid pMR12. Cells of *E. coli* strain DH1 $\lambda^+$  were used for primary transformation, and strain M5248 ( $\lambda$  cl857) was used for protein expression, although any comparable pair of *E. coli* strains carrying the cl $^+$  and cl857 alleles could be utilized. Alternatively, any rec $^+$  cl $^+$  strain could be induced by chemical agents such as nalidixic acid to produce the polymerase.

#### Purification of Polymerase

M5248 containing plasmid pMR12 was grown in one liter of LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl), preferably 2X LB medium, containing 100 mg/ml ampicillin at 30°C. When the OD<sub>600</sub> reached 1.0, the culture was induced at 42°C for 1.5 hours. The cultures were then cooled to <20°C and the cells harvested by centrifugation in a Sorvall RC-3B centrifuge at 5000 rpm at 4°C for 15 to 30 minutes. Harvested cells were stored at -80°C.

The cell pellet was resuspended in 25 ml pre-warmed lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM MgCl<sub>2</sub>, 16 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM EDTA, 0.1%, preferably 0.2% Tween 20, 0.1%, preferably 0.2% NP40). Preferably, the lysis buffer contains 300 mM NaCl. Resuspended cells were incubated at 75 - 85°C for 10-20 minutes, sonicated for 1 minute, and

cleared by centrifugation. The cleared lysate was passed through a 300 ml column of diethylaminoethyl cellulose (Whatman DE 52) equilibrated in buffer A (50mM Tris-HCl pH 8.0, 1mM EDTA, 0.1% Tween 20, 0.1% NP40) containing 100mM, preferably 300 mM NaCl. Fractions were assayed for polymerase activity, and those demonstrating peak polymerase activity were pooled, diluted to 50 mM NaCl with Buffer A, and loaded onto a heparin sepharose column (20 ml) equilibrated with 50 mM NaCl in buffer A. The polymerase was eluted from the column with a linear salt gradient from 50 mM to 700mM NaCl in buffer A. Fractions were assayed for polymerase activity, and those demonstrating peak activity were pooled and dialyzed against final buffer (20mM Tris-HCl pH8.5, 50 % (v/v) glycerol, 0.1mM EDTA, 0.5% Tween 20, 0.5% NP40, 1mM DTT, 100mM KCl). The purified protein is designated FY7. The amino acid sequence (and DNA sequence encoding therefor) are presented in Figure 1.

#### Bacterial Strains

*E. coli* strains: DH $\lambda$ <sup>+</sup> [gyrA96, recA1, relA1, endA1, thi-1, hsdR17, supE44,  $\lambda$ <sup>+</sup>]; M5248 [ $\lambda$  (bio275, cl857, cIII+, N+,  $\lambda$  (H1))].

#### PCR

Plasmid DNA from *E. coli* DH $\lambda$ <sup>+</sup> (pMR10) was prepared by SDS alkaline lysis method (Sambrook et al., Molecular Cloning 2<sup>nd</sup> Ed. Cold Spring Harbor Press, 1989). Reaction conditions were as follows: 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.001% gelatin, 1uM each primer, 2.5U Taq polymerase, per 100  $\mu$ l reaction. Cycling conditions were 94°C 2 minutes, then 35 cycles of 94°C 30s, 55°C 30s, 72°C 3 minutes, followed by 72°C for 7 minutes.

#### Example 1 Formulation of the enzyme in Mn conditions

In the following "pre-mix" protocol, all the reagents are contained in two solutions; reagent mix A and reagent mix B.

#### Reagent Mix A

The following reagents were combined to make 10 ml of reagent mix A:

2.5 ml 1 M HEPES N-(2-hydroxyethyl) piperazine-N'-(3-propanesulfonic acid), pH 8.0  
500  $\mu$ l 1 M tartaric acid, pH 8.0  
50,000 units FY7 DNA polymerase  
1 unit *Thermoplasma acidophilum* inorganic pyrophosphatase  
100  $\mu$ l 100 mM dATP  
100  $\mu$ l 100 mM dTTP

100 µl 100 mM dCTP  
500 µl 100 mM dITP  
9.375 µl 100 µM C-7-propargylamino-4-rhodamine-6-G-ddATP  
90 µl 100 µM C-5-propargylamino-4-rhodamine-X-ddCTP  
6.75 µl 100 µM C-7-propargylamino-4-rhodamine-110-ddGTP  
165 µl 100 µM C-5-propargylamino-4-tetramethylrhodamine-ddUTP  
10 µl 50 mM EDTA  
1 ml glycerol

The volume was made up to 10,000 µl with deionized H<sub>2</sub>O.

Reagent Mix B

The following reagents were combined to make 10 ml of reagent mix B:

10 µl 1M MES 2-(N-morpholino)ethanesulfonic acid, pH 6.0  
200 µl 1M MgCl<sub>2</sub>  
75 µl 1M MnSO<sub>4</sub>

The volume was made up to 10,000 µl with deionized H<sub>2</sub>O.

Example 2: Use of the formulation from Example 1

Two (2) µl reagent mix A, 2 µl reagent mix B, 200 ng M13mp18 DNA, 5 pmole of primer (M13 - 40 Forward 5'-GTTTCCCAGTCACGACGTTGTA), and deionized water to a total volume of 20 µl were mixed together and subjected to 25 cycles of (95 °C 30 seconds, 60 °C 1 minute) in a thermal cycler. After cycling, 4 µl of a solution which contained 1.5 M sodium acetate, 250 mM EDTA was added. The solution was mixed and 4 volumes (100 µl) of ethanol added. The DNA was precipitated by incubation on ice for 15-20 minutes followed by centrifugation. The supernatant was removed and the pellet was washed with 70% ethanol, dried and resuspended in 4 µl of formamide containing loading dye. The resuspended DNA was then run on an automated fluorescent DNA sequencing apparatus (ABI model 377 instrument). The print out from the machine of the DNA sequence is shown as Figure 2.

Example 3 Formulation of the enzyme in Mg conditions

In the following "pre-mix" protocol, all the reagents are contained in one solution.

Sequencing premix

The following reagents were combined to make 800 µl of Sequencing premix  
200 µl of 500 mM Tris-HCl pH 9.5, 20 mM MgCl<sub>2</sub>

100  $\mu$ l 40 units/ $\mu$ l FY7 DNA polymerase, 0.0008 units/ $\mu$ l *Thermoplasma acidophilum* inorganic pyrophosphatase

100  $\mu$ l 10 mM dITP, 2 mM dATP, 2 mM dTTP, 2 mM dCTP

100  $\mu$ l 0.125  $\mu$ M C-7-propargylamino-4-rhodamine-6-G-ddATP

100  $\mu$ l 1.2  $\mu$ M C-5-propargylamino-4-rhodamine-X-ddCTP

100  $\mu$ l 0.09  $\mu$ M C-7-propargylamino-4-rhodamine-110-ddGTP

100  $\mu$ l 2.2  $\mu$ M C-5-propargylamino-4-tetramethylrhodamine-ddUTP

Example 4 Use of the formulation from example 3

Four (4)  $\mu$ l of sequencing premix, 200 ng M13mp18 DNA, 5 pmole of primer (M13 - 40 Forward 5'- GTTTTCCCAGTCACGACGTTGTA), and deionized water to a total volume of 20  $\mu$ l were mixed together and subjected to 25 cycles of (95 °C 30 seconds, 60 °C 2 minutes) in a thermal cycler. After cycling, 7  $\mu$ l of 7.5 M ammonium acetate was added. The solution was mixed and 4 volumes (100  $\mu$ l) of ethanol added. The DNA was precipitated by incubation on ice for 15-20 minutes followed by centrifugation. The supernatant was removed and the pellet was washed with 70% ethanol, dried and resuspended in 4  $\mu$ l of formamide containing loading dye. The resuspended DNA was then run on an automated fluorescent DNA sequencing apparatus (ABI model 377 instrument). The print out from the machine of the DNA sequence is shown as Figure 3.

Example 5 Polymerase Activity versus Salt Concentration (KCl) for Thermo Sequenase™ enzyme and FY7 enzyme.

The percent of maximum polymerase activity was measured for Thermo Sequenase™ enzyme DNA polymerase and FY7 DNA polymerase under varying KCl concentrations. The results are depicted in Figure 4. The data indicate that FY7 has a much higher salt optimum as well as broader range of tolerance for salt in the reaction mixture than Thermo Sequenase™. The salt concentration which gives 50% activity is five-fold higher for FY7 than for Thermo Sequenase.

The effect of high salt concentrations on DNA sequencing ability in radioactively labeled DNA sequencing reactions was also examined. The results are presented in Figure 5. At KCl concentrations of 50mM or higher Thermo Sequenase™ polymerase performance degrades to levels at which usable data cannot be extracted. FY7 DNA polymerase, however, is able to give quite good sequencing data at concentrations of KCl of 100mM.

Example 6 Fluorescent Sequencing Salt Tolerance

These experiments examined the effect of the above-demonstrated polymerase activity in high salt concentrations on DNA sequencing ability in fluorescently labeled terminator DNA sequencing reactions. The results are presented in Figures 6-15.

Figures 6-10 show the effect of increasing salt concentration on the performance of Thermo Sequenase. At concentrations as low as 25mM data quality is affected with the read length being decreased from at least 600 bases to about 450 bases. At 50mM salt the read length is further decreased to about 350 bases, 75mM to about 250 bases and at 100mM the read length is negligible.

Figures 11-15 show the effect of increasing salt concentration on the performance of FY7 DNA polymerase. There is no detrimental effect on performance to at least 75mM KCl and only a slight decrease in data quality at 100mM KCl.

As it is recognized that some types of DNA preparations may be contaminated with salt (which is detrimental to DNA sequencing data quality), the use of FY7 DNA polymerase allows for a more robust sequencing reaction over a broader range of template conditions.

#### Example 7 Polymerase Processivity

The processivity (number of nucleotides incorporated per DNA polymerase binding event) has been measured, for different DNA sequencing polymerases. The results are presented in Figure 16. Thermo Sequenase DNA polymerase has a processivity of only ~4 nucleotides per binding event. AmpliTaq FS DNA polymerase has a processivity of ~15 nucleotides per binding event. FY7 DNA polymerase has a processivity more than seven-fold greater than Thermo Sequenase DNA polymerase and ~two-fold greater than AmpliTaq FS DNA polymerase at ~30 nucleotides per binding event.

#### Example 8 Polymerase Extension with dITP at 72 °C

The series examined improved read length obtained when using FY7 polymerase versus Thermo Sequenase DNA polymerase in radioactively labeled sequencing reactions incorporating the dGTP (Guanosine triphosphate) analog dITP (Inosine triphosphate) at 72 °C. The results are presented in Figure 17. FY7 is able to incorporate >50-100 more nucleotides under standard  $^{33}\text{P}[\alpha\text{-dATP}]$  sequencing conditions than Thermo Sequenase.

#### Example 9 Effect of Extension Step Time on Length of Read

These series of experiments examined the effect of increasing extension step time of the read length and data quality of Thermo Sequenase and FY7 DNA polymerases in fluorescently labeled terminator DNA sequencing reactions. The results are presented in Figures 18-27.

Figures 18-22 show the effect of increasing extension step time on the read length and data quality produced by Thermo Sequenase DNA polymerase. This data shows that a minimum of a two

minutes extension step is required by Thermo Sequenase in order to achieve a quality read of at least 600 bases. Signal strength generally increases to a maximum at a four minute extension (the time specified in the commercial product utilizing this enzyme and method).

Figures 23-27 show the effect of increasing extension step time on the read length and data quality produced by FY7 DNA polymerase. This data shows that a minimum of a 30 second extension step is required by FY7 in order to achieve a quality read of at least 600 bases. Signal strengths plateau at about one minute extension time. The FY7 DNA polymerase can produce data of equivalent quality to Thermo Sequenase in one-quarter to one-half the time of extension reaction.

Although the above examples describe various embodiments of the invention in detail, many variations will be apparent to those of ordinary skill in the art. Accordingly, the above examples are intended for illustration purposes and should not be used in any way to restrict the scope of the appended claims.

## CLAIMS

What is claimed is:

1. A purified recombinant thermostable DNA polymerase comprising the amino acid sequence set forth in Figure 1.
2. A purified recombinant thermostable DNA polymerase which exhibits at least about 80% activity at salt concentrations of 50 mM and greater.
3. A purified recombinant thermostable DNA polymerase which exhibits at least about 70% activity at salt concentrations of 25 mM and greater.
4. A purified recombinant thermostable DNA polymerase having a processivity of about 30 nucleotides per binding event.
5. An isolated nucleic acid that encodes a thermostable DNA polymerase, wherein said nucleic acid consists of the nucleotide sequence set forth in Figure 1
6. A recombinant DNA vector that comprises the nucleic acid of Claim 3.
7. The recombinant DNA sequence of Claim 4 comprising the plasmid pMR10.
8. A recombinant host cell transformed with the vector of Claim 5.
9. The recombinant host cell of Claim 6 that is *E. coli*.
10. The recombinant host cell of Claim 7 which is *E. coli* carrying the cl<sup>+</sup> and cl857 alleles.
11. The recombinant host cell of Claim 7 selected from the group consisting of DH1λ<sup>+</sup> [gyrA96, recA1, relA1, endA1, thi-1, hsdR17, supE44, λ<sup>+</sup>] and M5248 [λ (bio275, cl857, cIII+, N+, λ (H1))].
12. Method of sequencing DNA comprising the step of generating chain terminated fragments from the DNA template to be sequenced with the DNA polymerase of Claim 1 in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.

**13. A kit for sequencing DNA comprising the DNA polymerase of Claim 1.**

31/11  
 /1 ATG GAA GCG ATG CTG CCG CTG TTC GAA CCC AAA GGC CGT GTC CTC CTG GTG GCC GGC CAC  
 M E A M L P L F E P K G R V L L V A G H  
 91/31  
 61/21 CAC CTG GCC TAC CGC ACC TTC TTC GCC CTG AAG GGC CTC ACC ACG AGC CGG GGC GAA CCG  
 H L A Y R T F F A L K G L T T S R G E P  
 151/51  
 121/41 GTG CAG GCG GTC TAC GGC TTC GCC AAG AGC CTC CTC AAG GCC CTG AAG GAG GAC GGG TAC  
 V Q A V Y G F A K S L L K A L K E D G Y  
 211/71  
 181/61 AAG GCC GTC TTC GTG GTC TTT GAC GCC AAG GCC CCC TCC TTC CGC CAC GAG GCC TAC GAG  
 K A V F V V F D A K A P S F R H E A Y E  
 271/91  
 241/81 GCC TAC AAG GCG GGG AGG GCC CCG ACC CCC GAG GAC TTC CCC CGG CAG CTC GCC CTC ATC  
 A Y K A G R A P T P E D F P R Q L A L I  
 301/101 331/111  
 AAG GAG CTG GTG GAC CTC CTG GGG TTT ACC CGC CTC GAG GTC CCC GGC TAC GAG GCG GAC  
 K E L V D L L G F T R L E V P G Y E A D  
 361/121 391/131  
 GAC GTT CTC GCC ACC CTG GCC AAG AAG GCG GAA AAG GAG GGG TAC GAG GTG CGC ATC CTC  
 D V L A T L A K K A E K E G Y E V R I L  
 421/141 451/151  
 ACC GCC GAC CGC GAC CTC TAC CAA CTC GTC TCC GAC CGC GTC GCC GTC CTC CAC CCC GAG  
 T A D R D L Y Q L V S D R V A V L H P E  
 511/171  
 481/161 GGC CAC CTC ATC ACC CCG GAG TGG CTT TGG GAG AAG TAC GGC CTC AGG CCG GAG CAG TGG  
 G H L I T P E W L W E K Y G L R P E Q W  
 571/191  
 541/181 GTG GAC TTC CGC GCC CTC GTG GGG GAC CCC TCC GAC AAC CTC CCC GGG GTC AAG GGC ATC  
 V D F R A L V G D P S D N I P G V K G I  
 601/201 631/211  
 GGG GAG AAG ACC GCC CTC AAG CTC CTC AAG GAG TGG GGA AGC CTG GAA AAC CTC CTC AAG  
 G E K T A L K L L K E W G S L E N L L K  
 661/221 691/231  
 AAC CTG GAC CGG GTA AAG CCA GAA AAC GTC CGG GAG AAG ATC AAG GCC CAC CTG GAA GAC  
 N L D R V K P E N V R E K I K A H L E D  
 721/241 751/251  
 CTC AGG CTC TCC TTG GAG CTC TCC CGG GTG CGC ACC GAC CTC CCC CTG GAG GTG GAC CTC  
 L R L S L E L S R V R T D L P L E V D L  
 781/261 811/271  
 GCC CAG GGG CGG GAG CCC GAC CGG GAG GGG CTT AGG GCC TTC CTG GAG AGG CTG GAA TTC  
 A Q G R E P D R E G L R A F L E R L E F  
 871/291  
 841/281 GGC AGC CTC CTC CAC GAG TTC GGC CTC CTG GAG GCC CCC GCC CCC CTG GAG GAG GCC CCC  
 G S L L H E F G L L E A P A P L E E A P  
 901/301 931/311  
 TGG CCC CCG CCG GAA GGG GCC TTC GTG GGC TTC GTC CTC TCC CGC CCC GAG CCC ATG TGG  
 W P P P E G A F V G F V L S R P E P M W  
 961/321 991/331  
 GCG GAG CTT AAA GCC CTG GCC GCC TGC AGG GAC GGC CGG GTG CAC CGG GCA GCA GAC CCC  
 A E L K A L A A C R D G R V H R A A D P  
 1021/341 1051/351  
 TTG GCG GGG CTA AAG GAC CTC AAG GAG GTC CGG GGC CTC CTC GCC AAG GAC CTC GCC GTC  
 L A G L K D L K E V R G L L A K D L A V  
 1081/361 1111/371  
 TTG GCC TCG AGG GAG GGG CTA GAC CTC GTG CCC GGG GAC GAC CCC ATG CTC CTC GCC TAC  
 L A S R E G L D L V P G D D P M L L A Y  
 1141/381 1171/391  
 CTC CTG GAC CCC TCC AAC ACC ACC CCC GAG GGG GTG GCG CGG CGC TAC GGG GGG GAG TGG  
 L L D P S N T T P E G V A R R Y G G E W  
 1201/401 1231/411

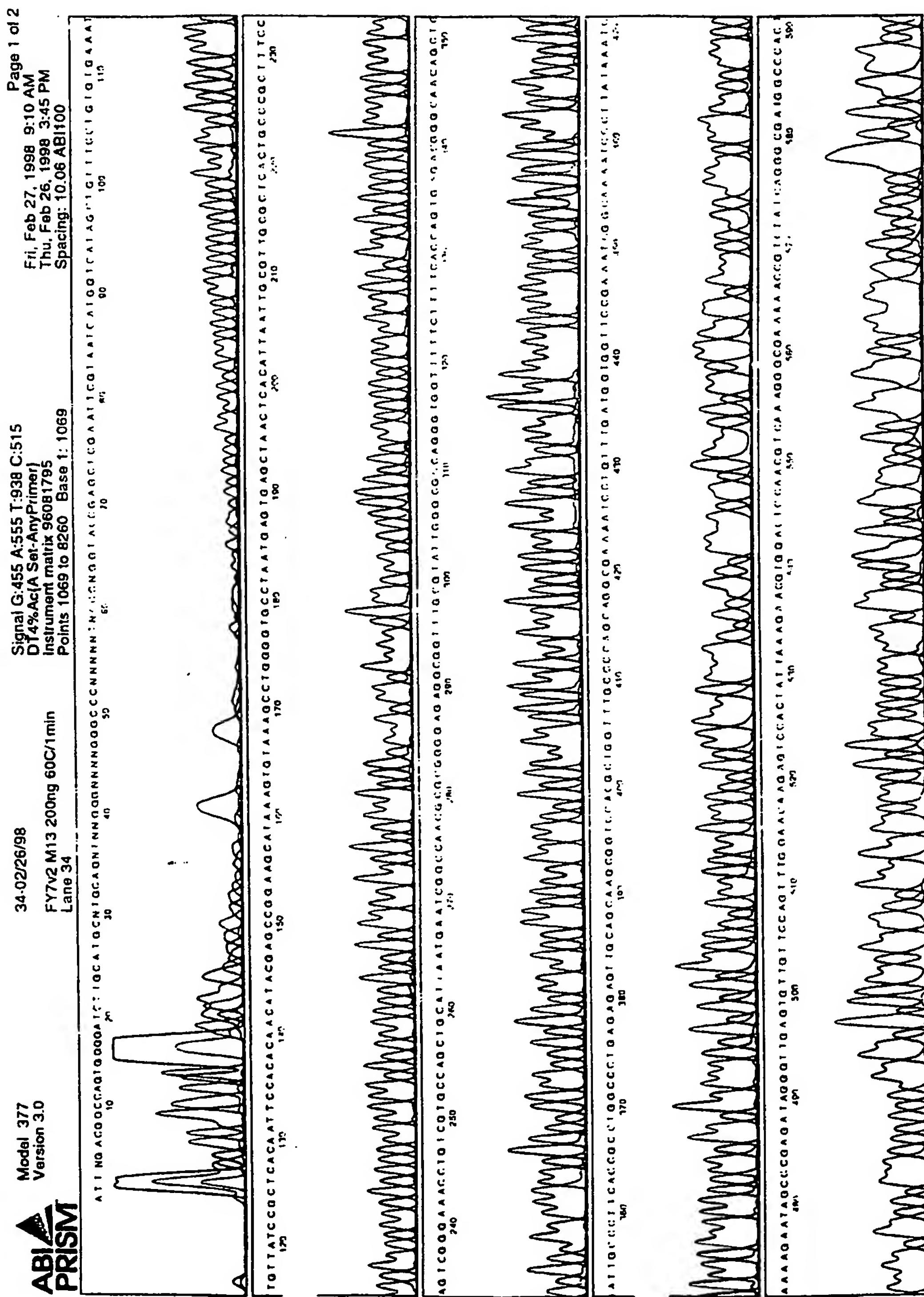
Figure 1

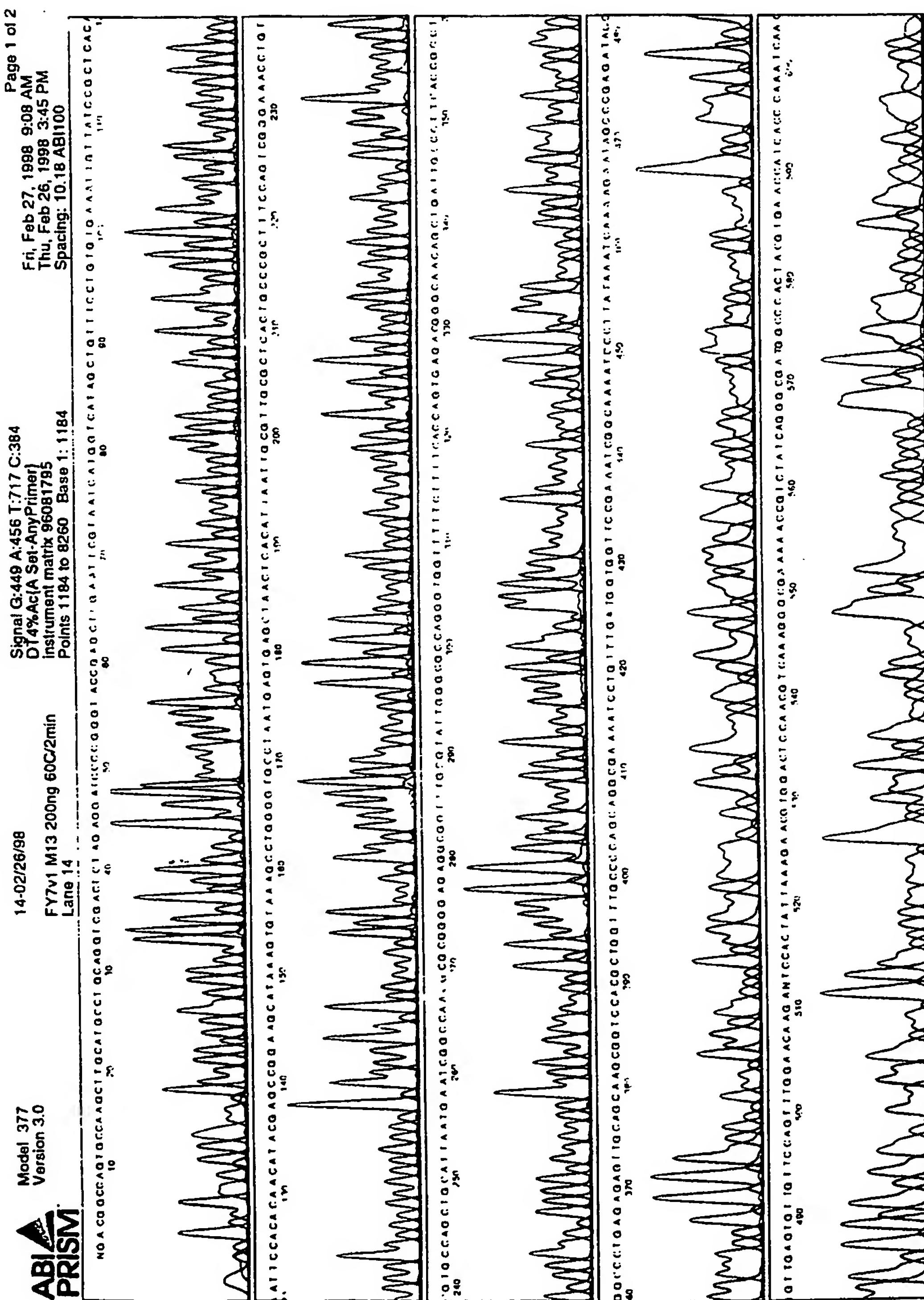
ACG GAG GAC GCC CAC CGG GCC CTC CTC TCG GAG AGG CTC CAT CGG AAC CTC CTT AAG  
 T E D A A H R A L L S E R L H R N L L K  
 1261/421 1291/431  
 CGC CTC GAG GGG GAG GAG AAG CTC CTT TGG CTC TAC CAC GAG GTG GAA AAG CCC CTC TCC  
 R L E G E E K L L W L Y H E V E K P L S  
 1321/441 1351/451  
 CGG GTC CTG GCC CAC ATG GAG ACC GGG GTA CGG CTG GAC GTG GCC TAC CTT CAG GCC  
 CGG GTC CTG GCC CAC ATG GAG ACC GGG GTA CGG CTG GAC GTG GCC TAC CTT CAG GCC  
 R V L A H M E A T G V R L D V A Y L Q A  
 1411/471  
 1381/461 CTT TCC CTG GAG CTT GCG GAG GAG ATC CGC CGC CTC GAG GAG GTC TTC CGC TTG GCG  
 L S L E L A E E I R R L E E E V F R L A  
 1441/481 1471/491  
 GGC CAC CCC TTC AAC CTC AAC TCC CGG GAC CAG CTG GAA AGG GTG CTC TTT GAC GAG CTT  
 GGC CAC CCC TTC AAC CTC AAC TCC CGG GAC CAG CTG GAA AGG GTG CTC TTT GAC GAG CTT  
 G H P F N L N S R D Q L E R V L F D E L  
 1501/501 1531/511  
 AGG CTT CCC GCC TTG GGG AAG ACG CAA AAG ACA GGC AAG CGC TCC ACC AGC GCC GCG GTG  
 R L P A L G K T Q K T G K R S T S A A V  
 1561/521 1591/531  
 CTG GAG GCC CTA CGG GAG GCC CAC CCC ATC GTG GAG AAG ATC CTC CAG CAC CGG GAG CTC  
 L E A L R E A H P I V E K I L Q H R E L  
 1621/541 1651/551  
 ACC AAG CTC AAG AAC ACC TAC GTG GAC CCC CTC CCA AGC CTC GTC CAC CCG AGG ACG GGC  
 T K L K N T Y V D ? L P S L V H P R T G  
 1681/561 1711/571  
 CGC CTC CAC ACC CGC TTC AAC CAG ACG GCC ACG GCC ACG GGC ACG GGG AGG CTT AGT AGC TCC GAC  
 R L H T R F N Q T A T A T G R L S S S D  
 1741/581 1771/591  
 CCC AAC CTG CAG AAC ATC CCC GTC CGC ACC CCC TTG GGC CAG AGG ATC CGC CGG GCC TTC  
 P N L Q N I P V R T P L G Q R I R R A F  
 1801/601 1831/611  
 GTG GCC GAG GCG GGT TGG GCG TTG GTG GCC CTG GAC TAT AGC CAG ATA GAG CTC CGC GTC  
 V A E A G W A L V A L D Y S Q I E L R V  
 1861/621 1891/631  
 CTC GCC CAC CTC TCC GGG GAC GAA AAC CTG ATC AGG GTC TTC CAG GAG GGG AAG GAC ATC  
 L A H L S G D E N I R V F Q E G K D I  
 1921/641 1951/651  
 CAC ACC CAG ACC GCA AGC TGG ATG TTC GGC GTC CCC CCG GAG GCC GTG GAC CCC CTG ATG  
 H T Q T A S W M F S V P P E A V D P L M  
 1981/661 2011/671  
 CGC CGG GCG GCC AAG ACG GTG AAC TAC GGC GTC CTC TAC GGC ATG TCC GCC CAT AGG CTC  
 R R A A K T V N Y G V L Y G M S A H R L  
 2041/681 2071/691  
 TCC CAG GAG CTA GCC ATC CCC TAC GAA GAA GCG GTG GCC TTT ATA GAG CGC TAC TTC CAA  
 S Q E L A I P Y E E A V A F I E R Y F Q  
 2101/701 2131/711  
 AGC TTC CCC AAG GTG CGG GCC TGG ATA GAA AAG ACC CTG GAG GAG GGG AGG AAG CGG GGC  
 S F P K V R A W I E K T L E E G R K R G  
 2161/721 2191/731  
 TAC GTG GAA ACC CTC TTC GGA AGA AGG CGC TAC GTG CCC GAC CTC AAC GCC CGG GTG AAG  
 Y V E T L F G R R R Y V P D L N A R V K  
 2221/741 2251/751  
 AGC GTC AGG GAG GCC GCG GAG CGC ATG GCC TTC AAC ATG CCC GTC CAG GGC ACC GCC GCC  
 S V R E A A E R M A F N M P V Q G T A A  
 2281/761 2311/771  
 GAC CTC ATG AAG CTC GCC ATG GTG AAG CTC TTC CCC CGC CTC CGG GAG ATG GGG GCC CGC  
 D L M K L A M V K I F P R L R E M G A R  
 2341/781 2371/791  
 ATG CTC CTC CAG GTC CAC GAC GAG CTC CTC CTG GAG GCC CCC CAA GCG CGG GCC GAG GAG  
 M L L Q V H D E I L L E A P Q A R A E E  
 2401/801 2431/811  
 GTG GCG GCT TTG GCC AAG GAG GCC ATG GAG AAG GCC TAT CCC CTC GCC GTG CCC CTG GAG  
 V A A L A K E A M E K A Y P L A V P L E

Figure 1 (continued)

2461/821 2491/831  
GTG GAG GTG GGG ATG GGG GAG GAC TGG CTT TCC GCC AAG GGT TAG  
V E V G M G E D W L S A K G \*

Figure 1 (continued)





**Polymerase Activity versus Salt Concentration (KCl)  
for Thermo Sequenase and FY7**

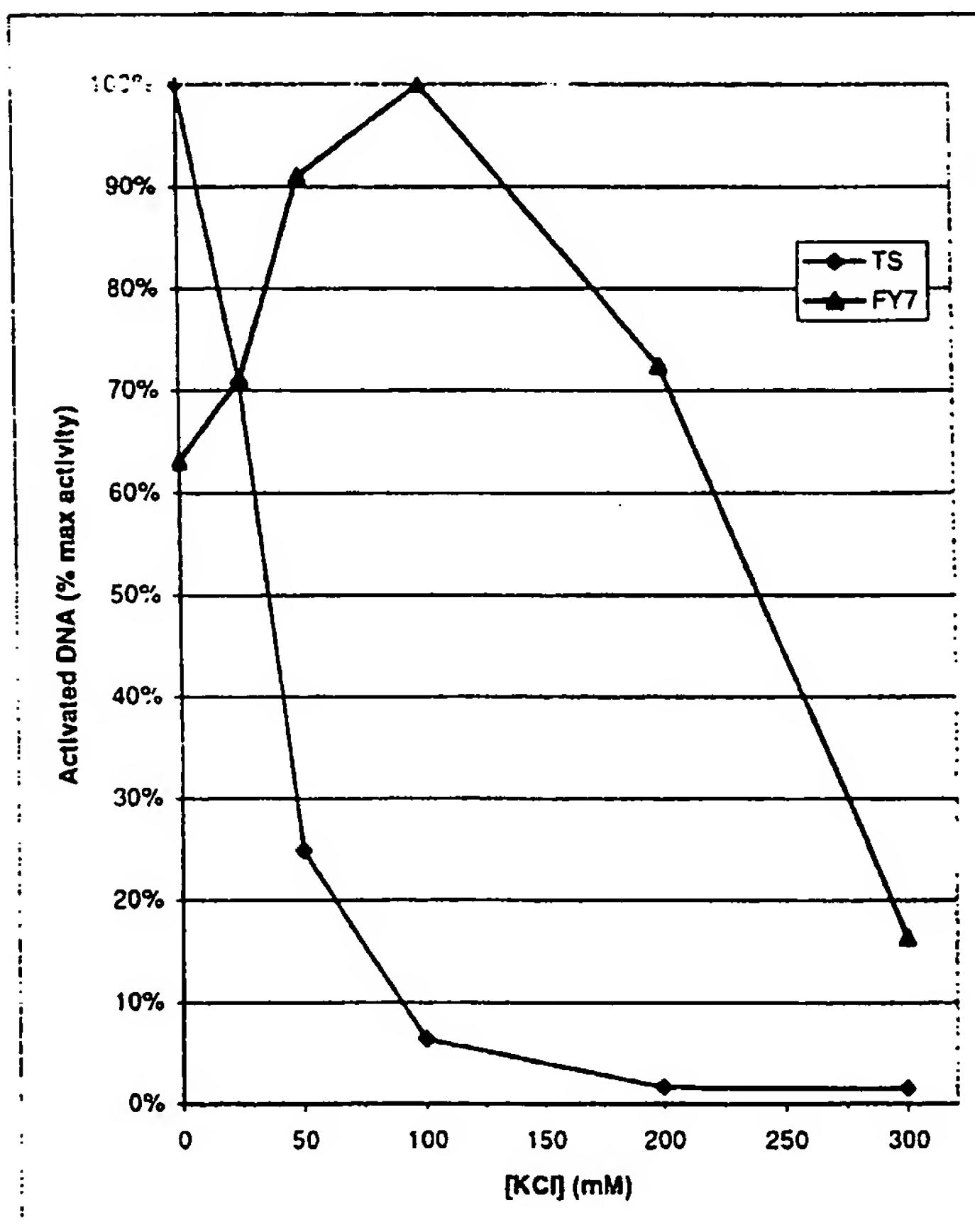


Figure 4

## SALT TOLERANCE (KCI) OF FY7 VS. THERMO SEQUENASE

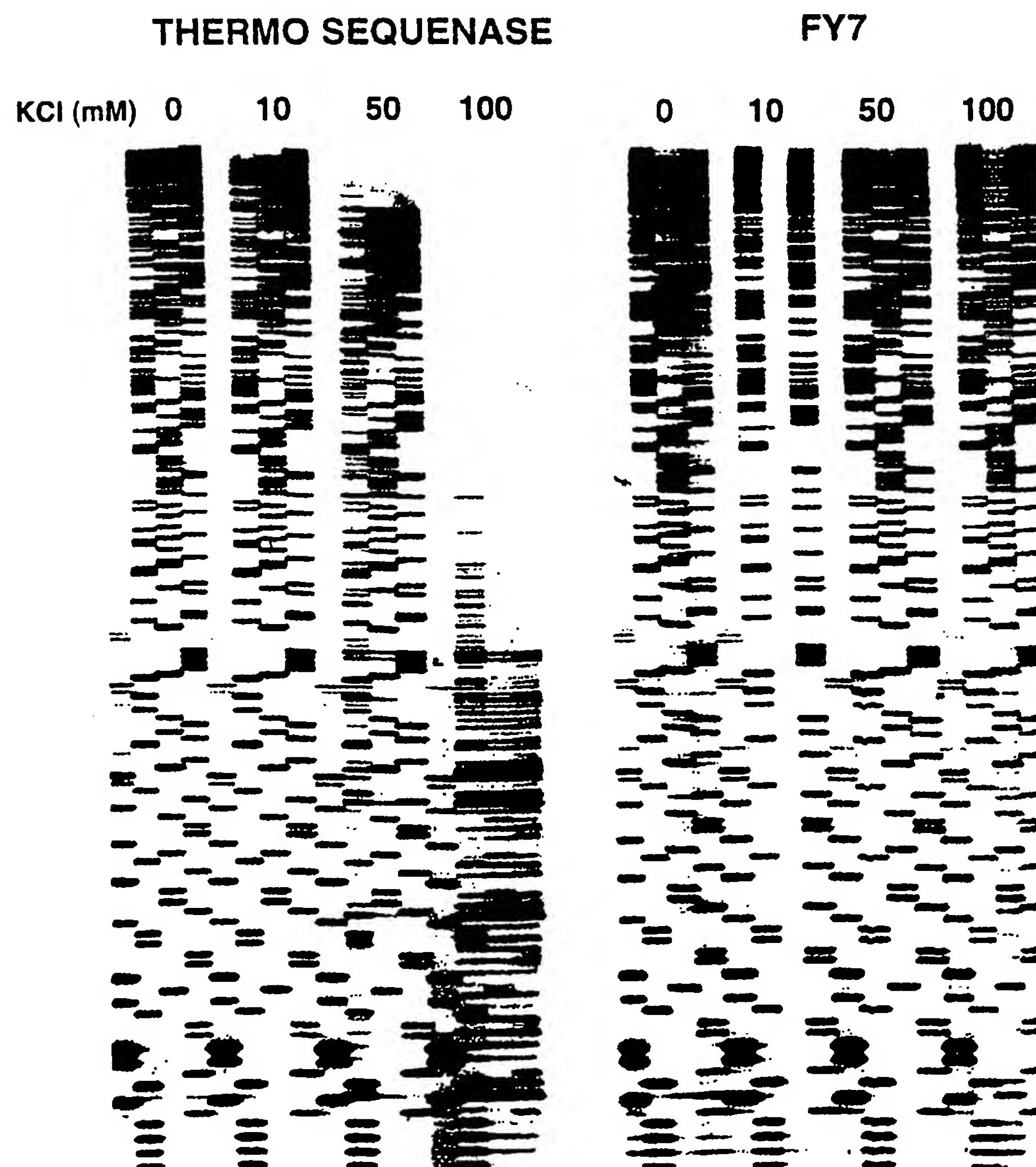
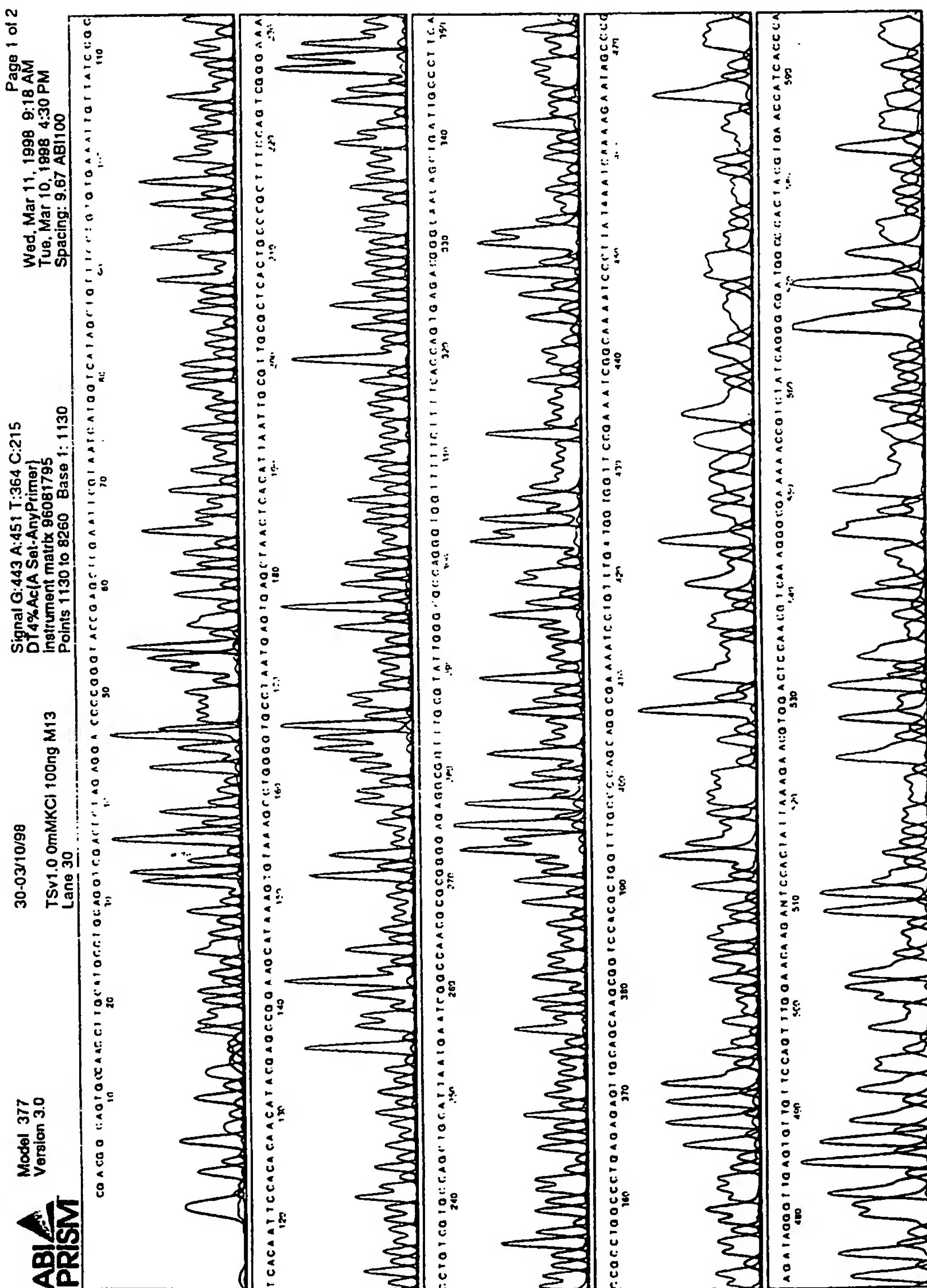


Figure 5



۶۰

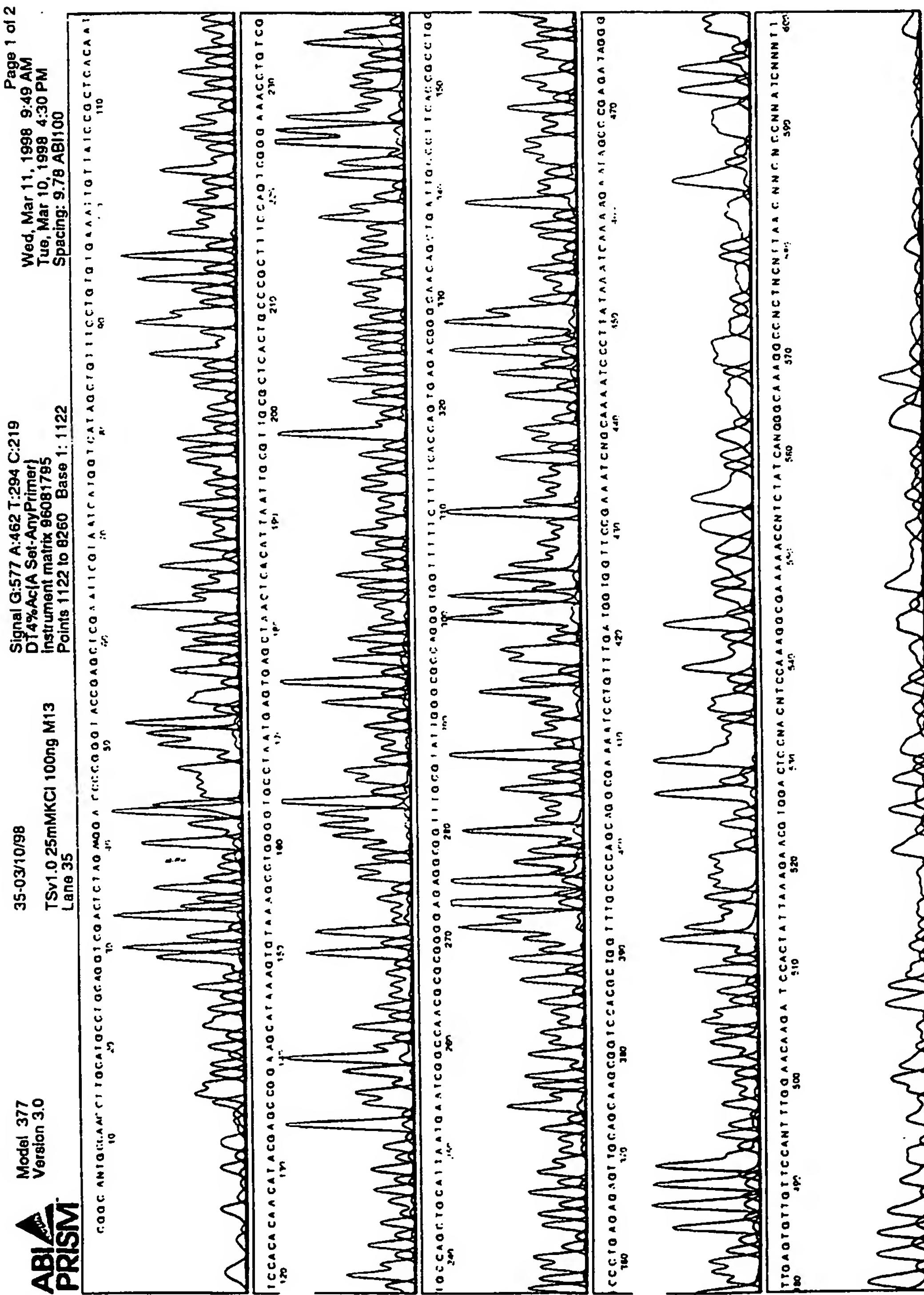
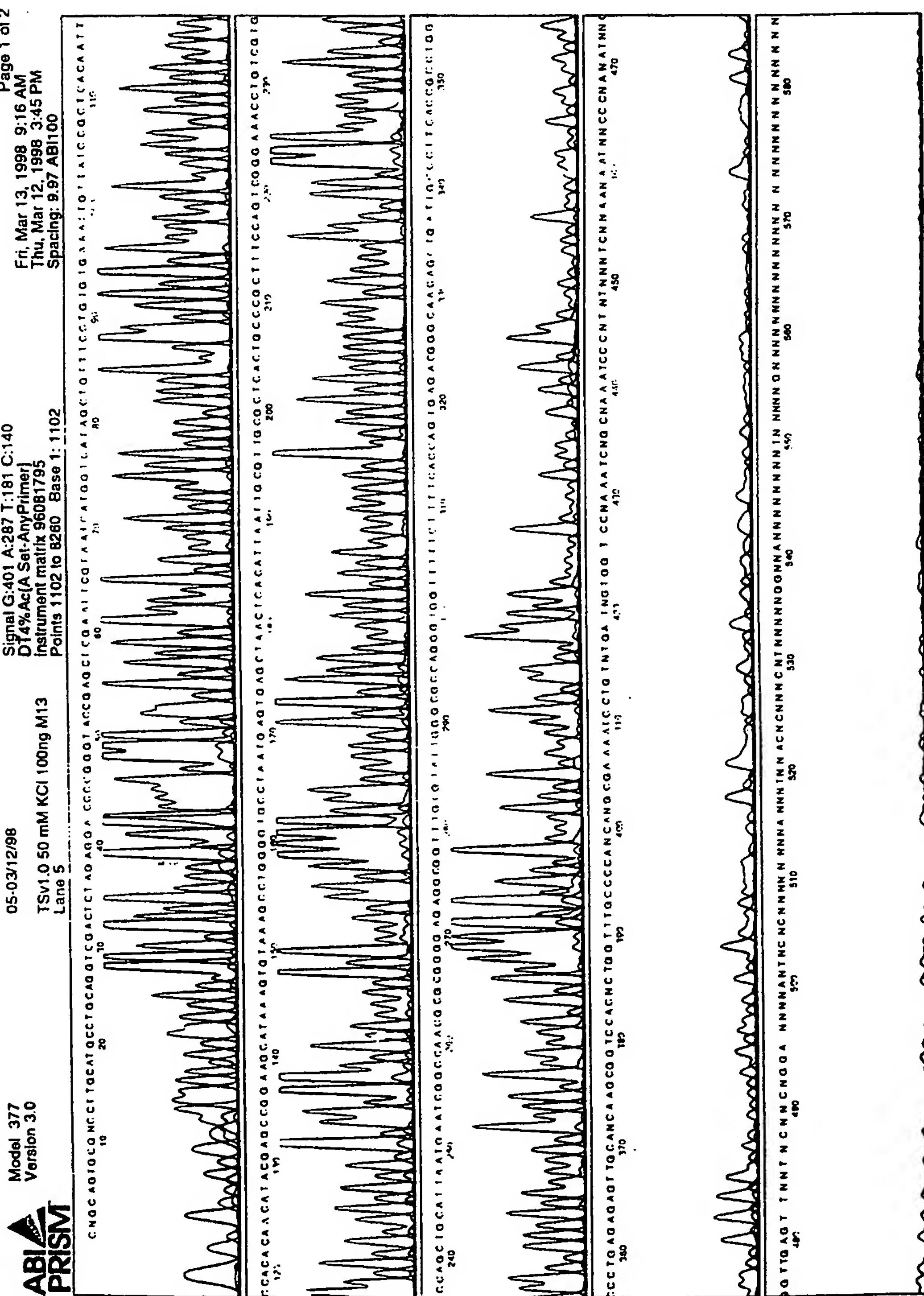


Figure 7



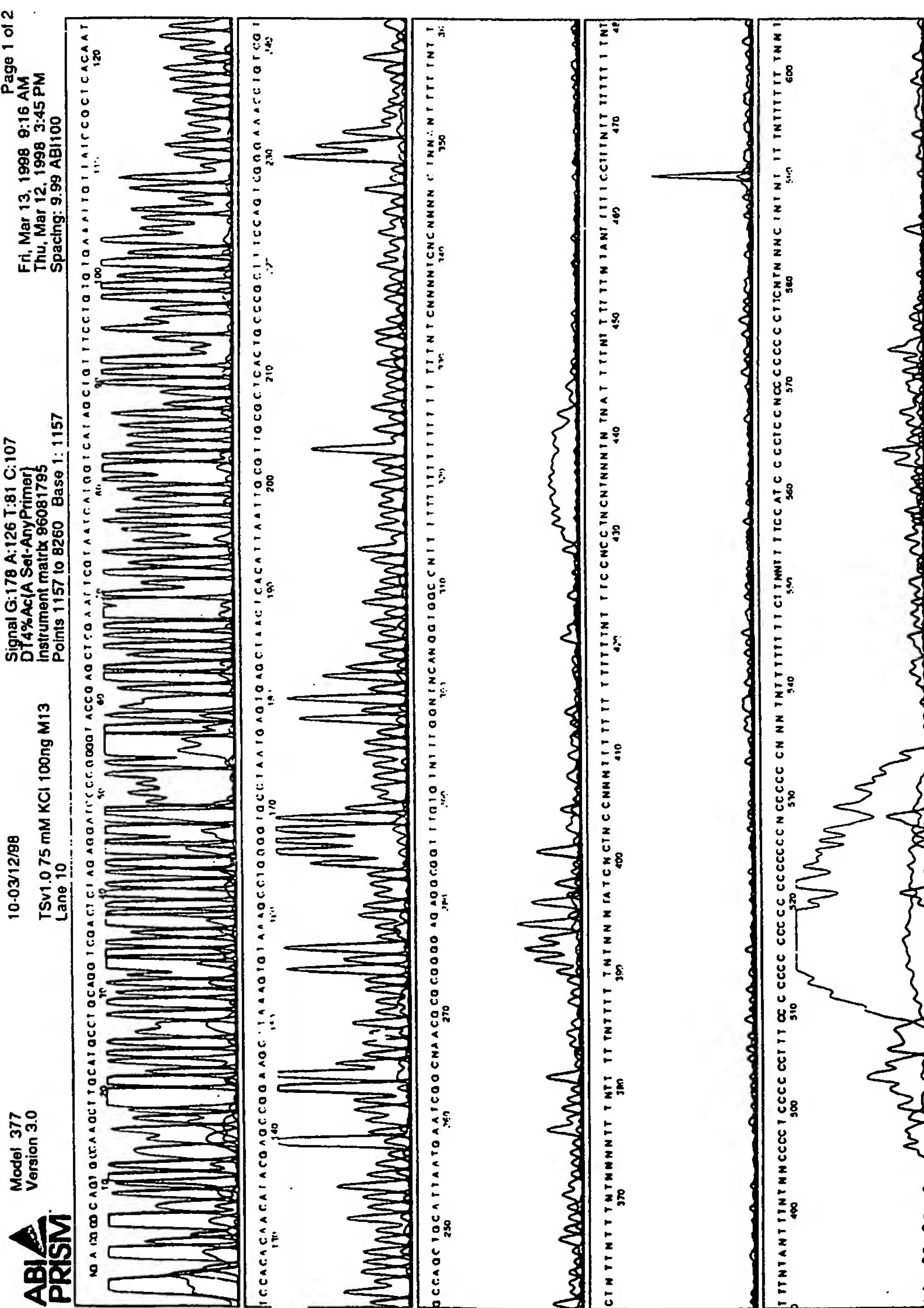


Figure 9

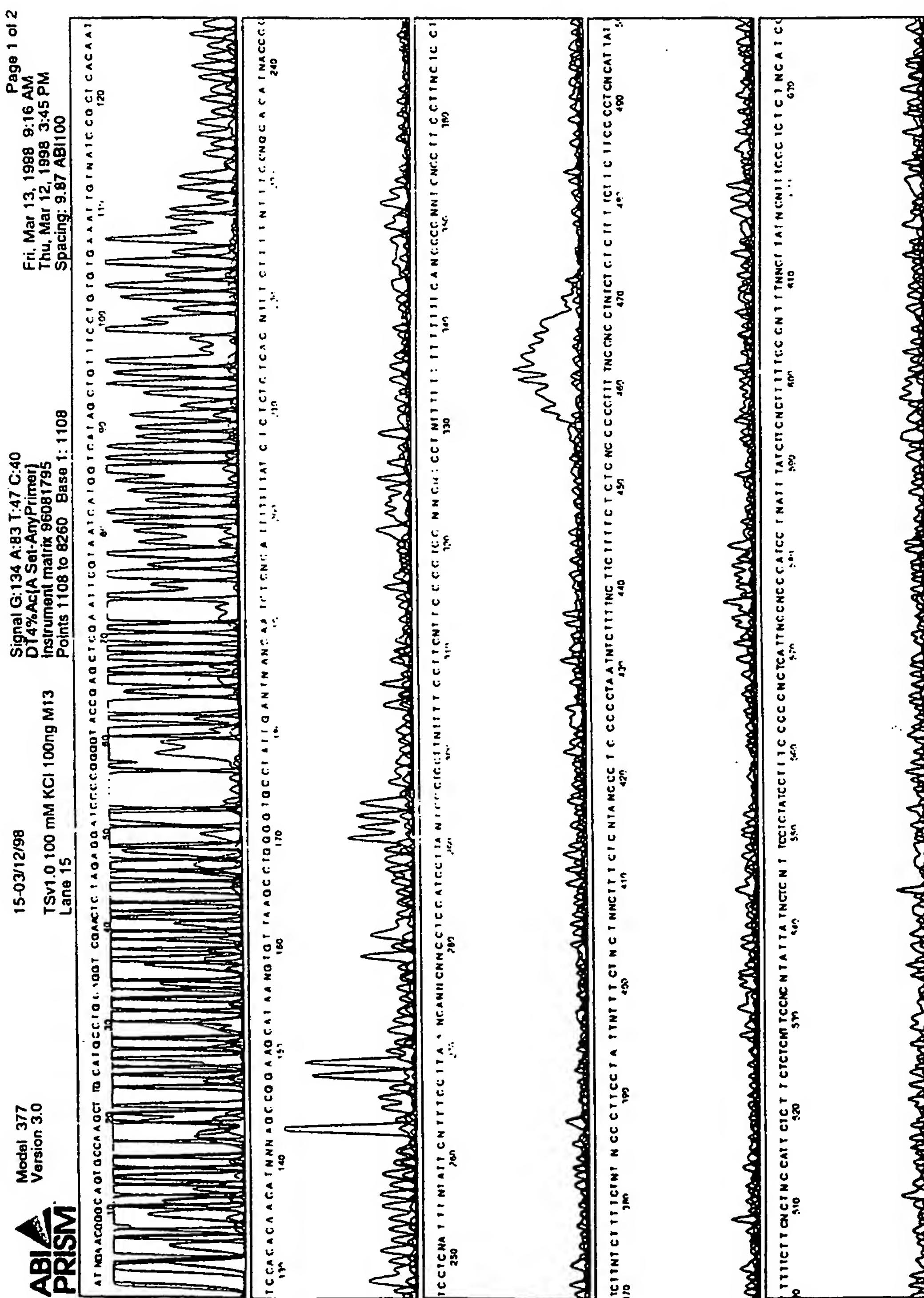
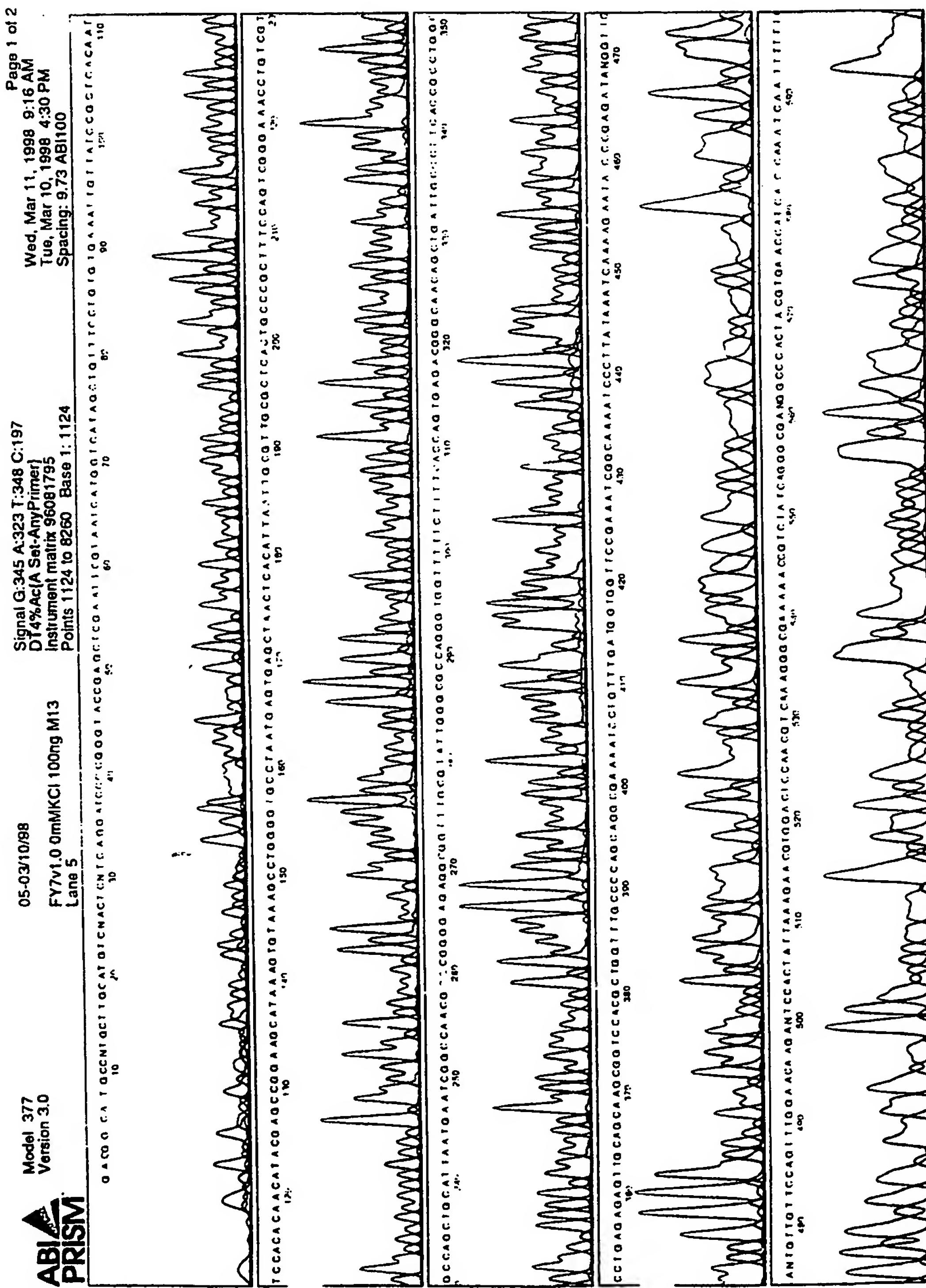


Figure 10



## Figure 1

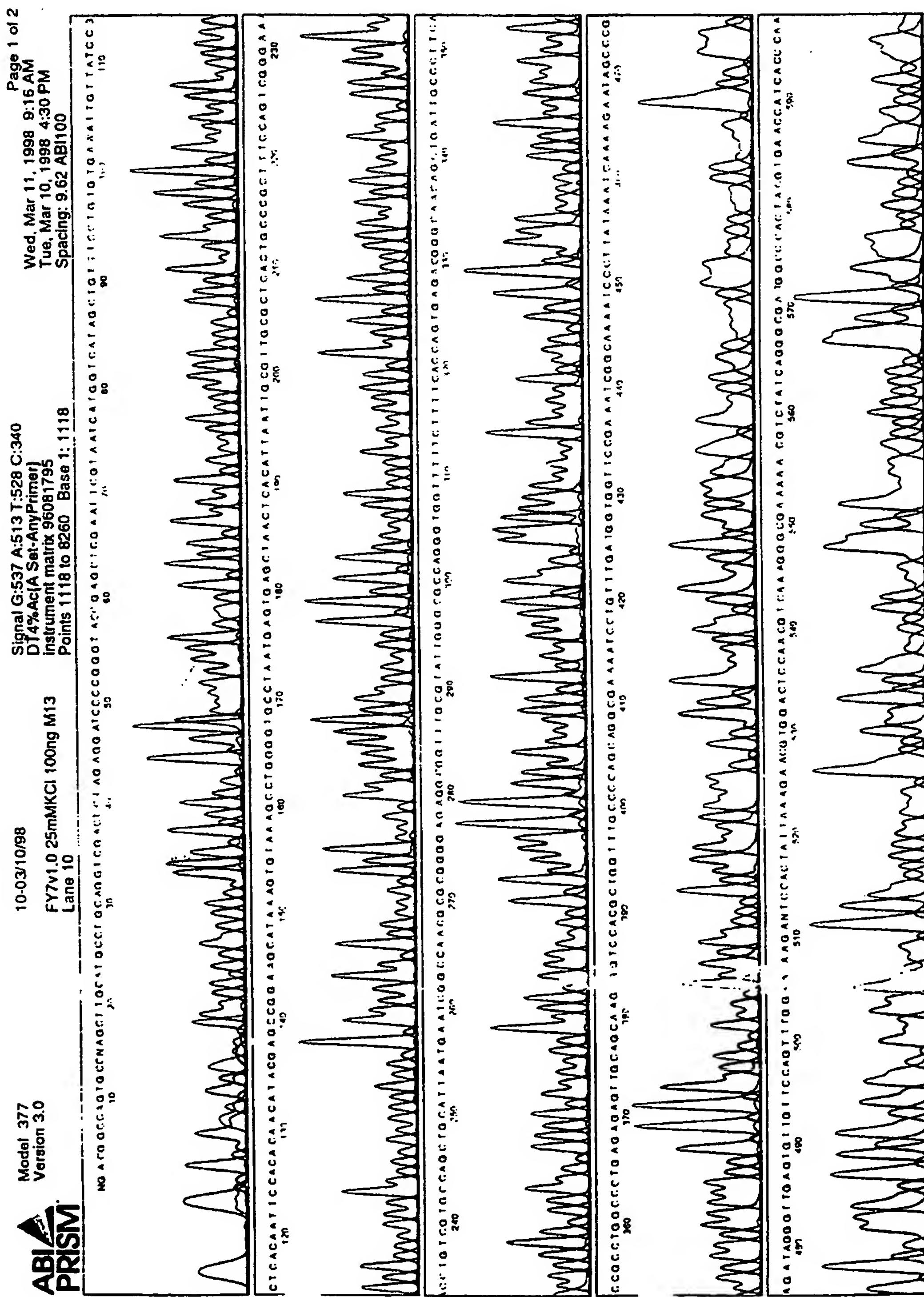


Figure 12

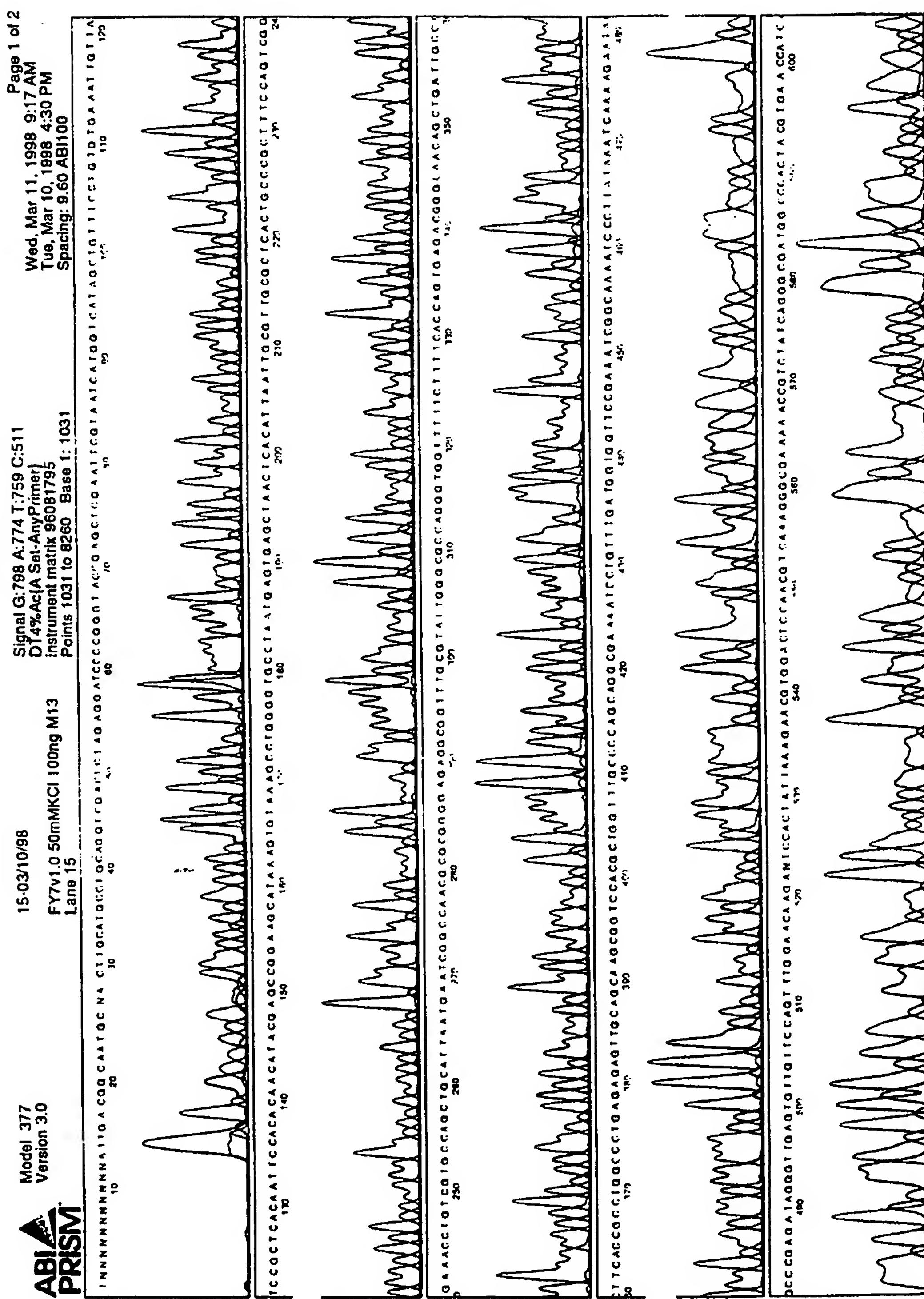
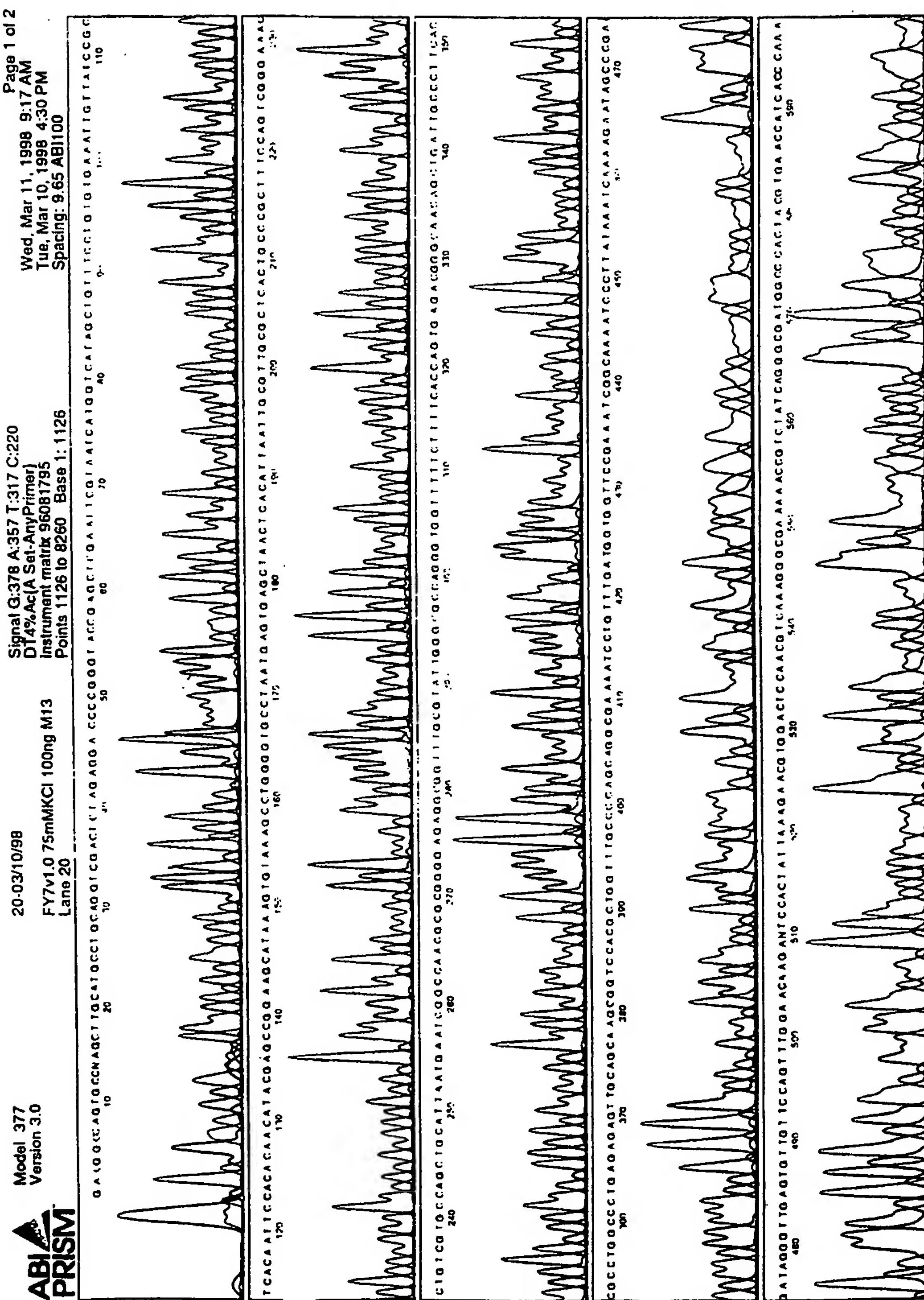


Figure 1.3

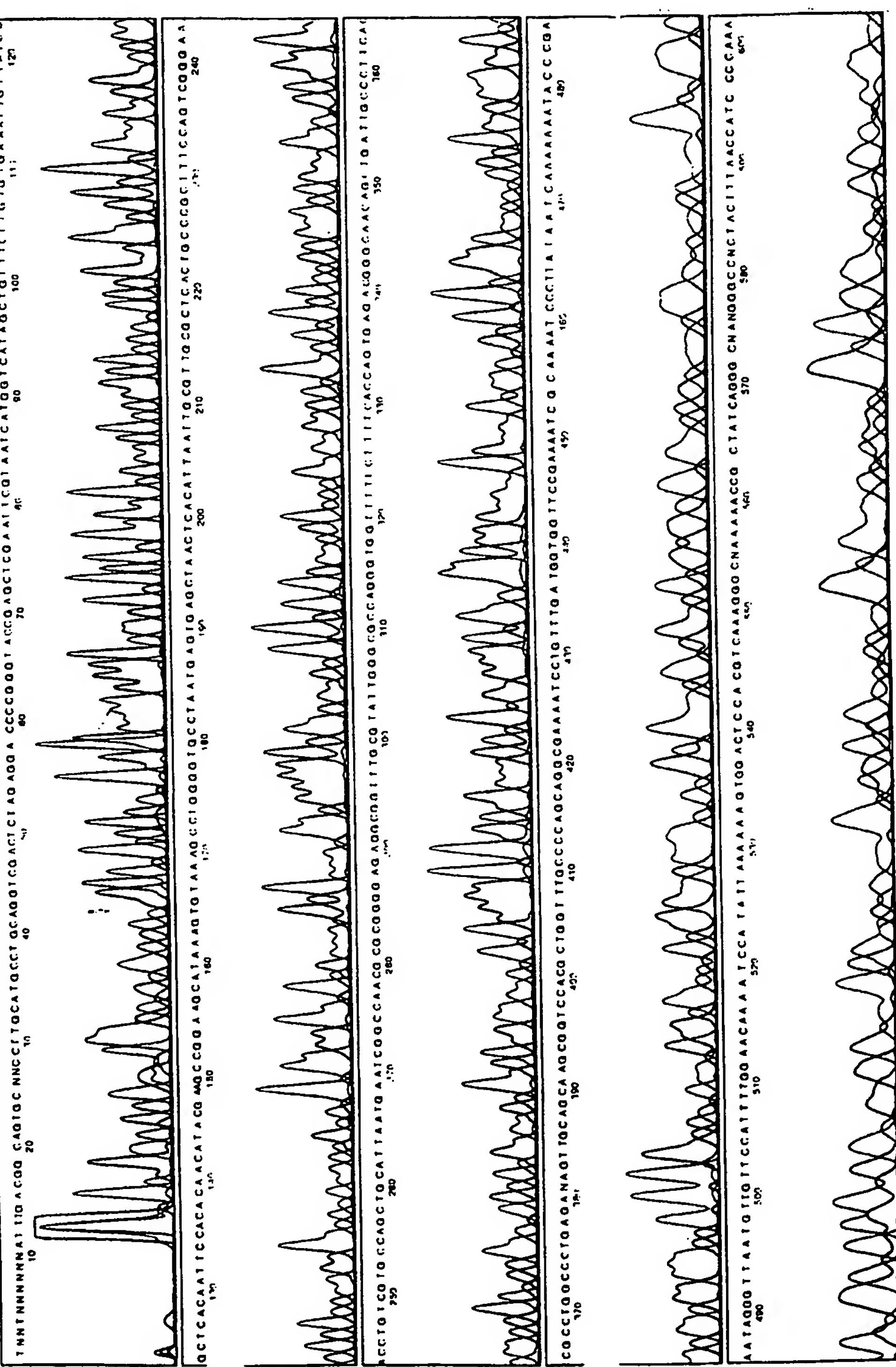


Liquide - H

Model 377  
Version 3.0

FY7v1.0 100  
Lane 25

Page 1 of 2  
Wed, Mar 11, 1998 9:17 AM  
Tue, Mar 10, 1998 4:30 PM  
Spacing: 9.72 AB/100



## Figure 15

Polymerase	Processivity <u>number of nucleotides</u>
Thermo Sequenase	4
AmpliTaq FS	15
FY7	30

Figure 16

## Polymerase Extension with dITP at 72°C

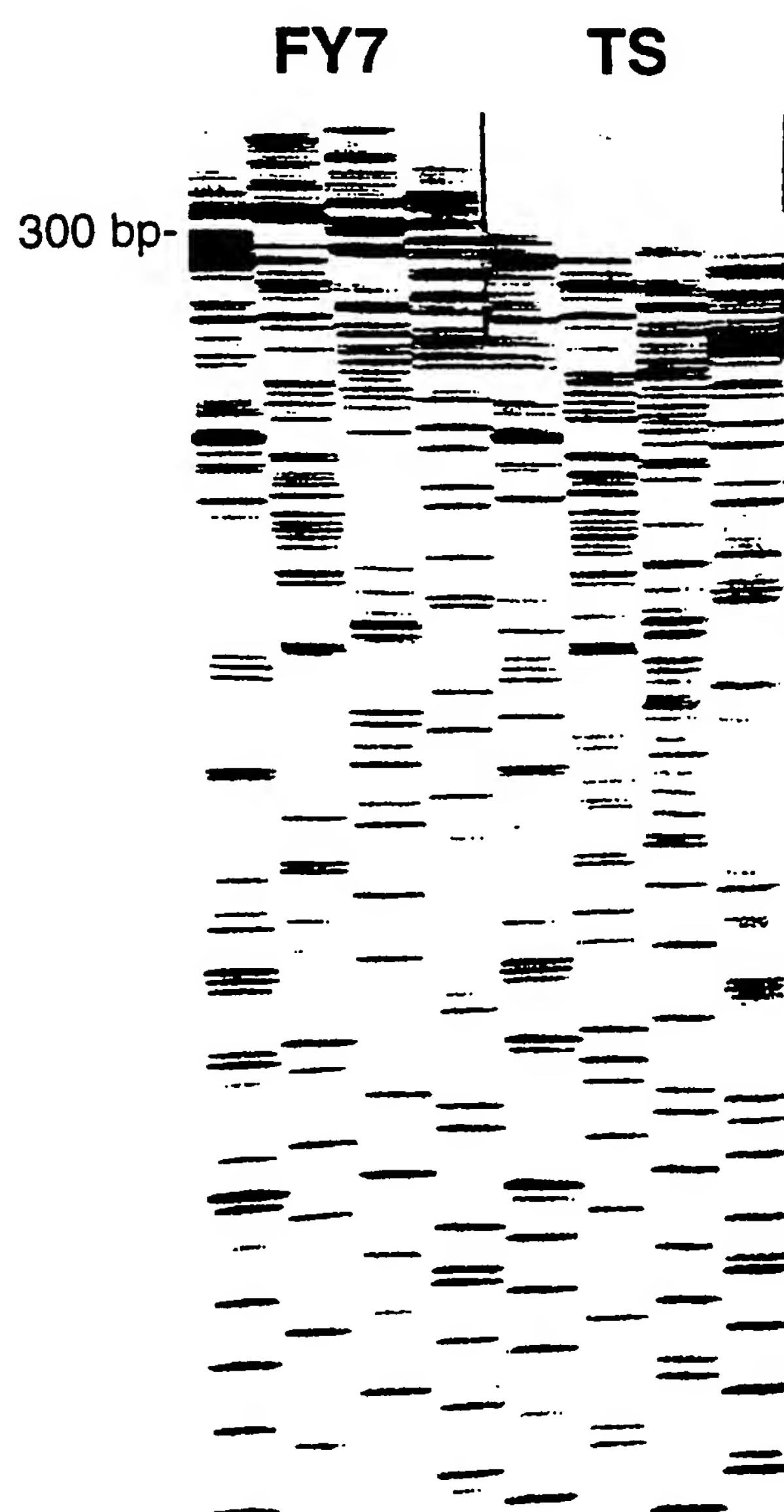


Figure 17

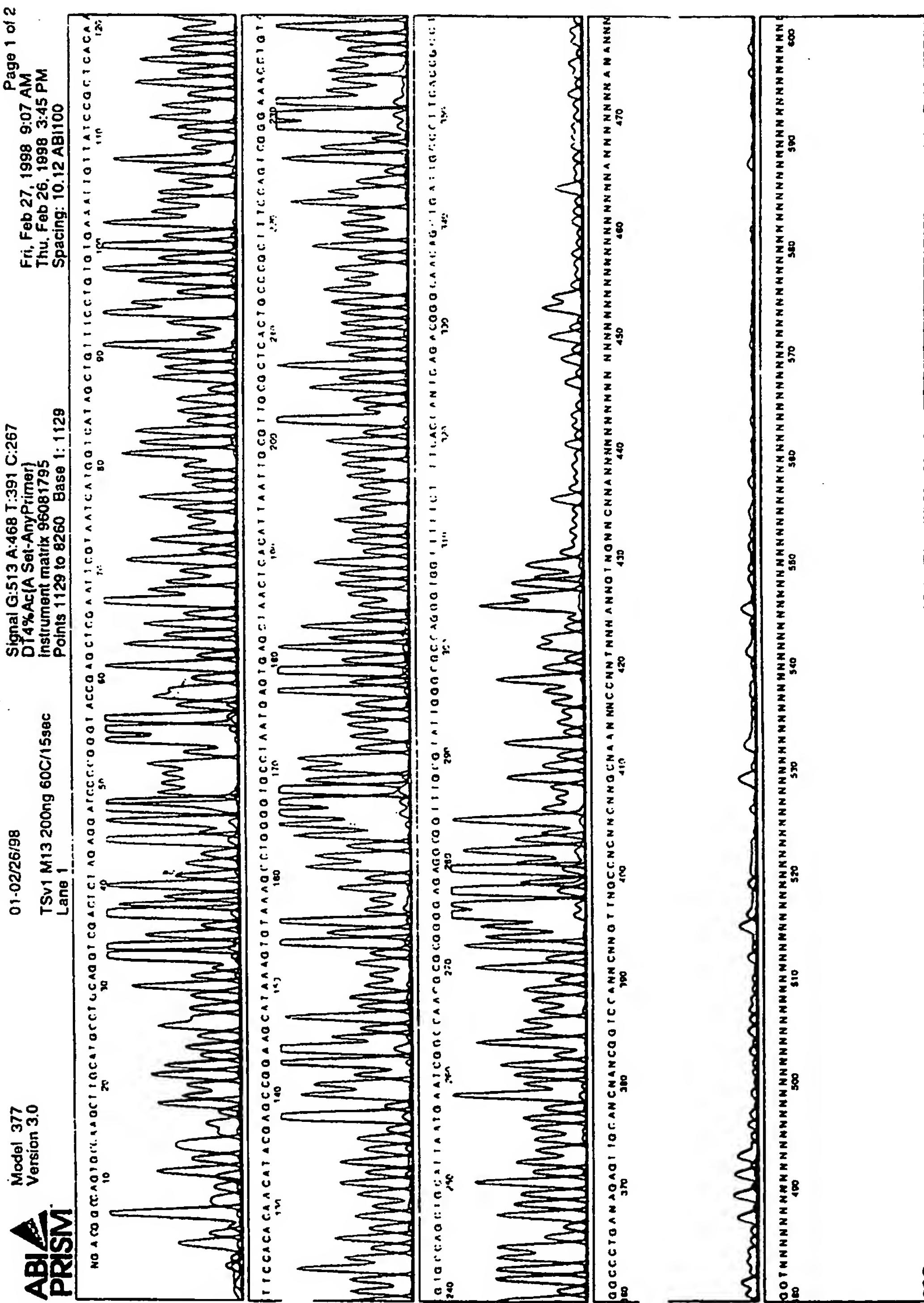


Figure 18

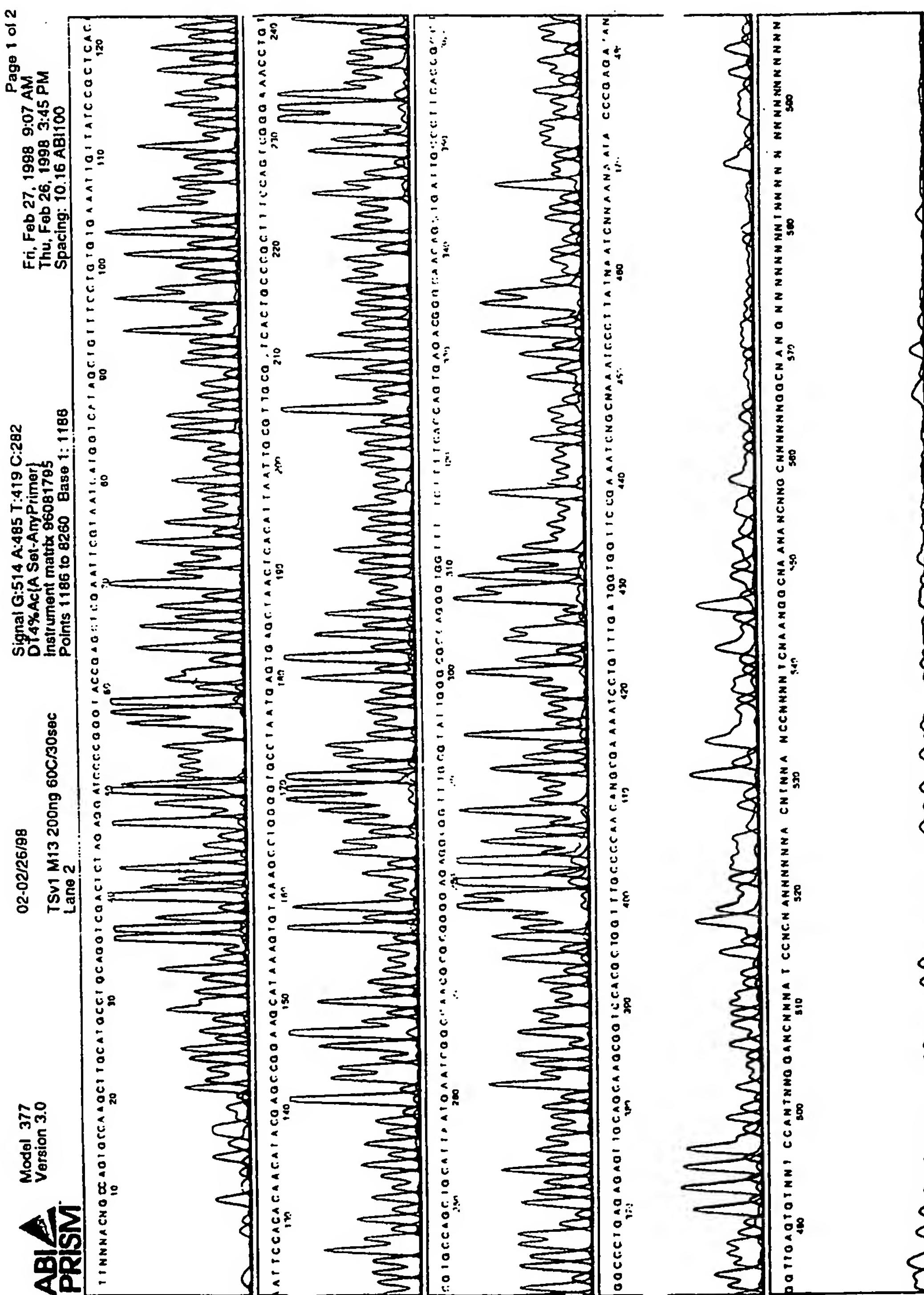


Figure 19

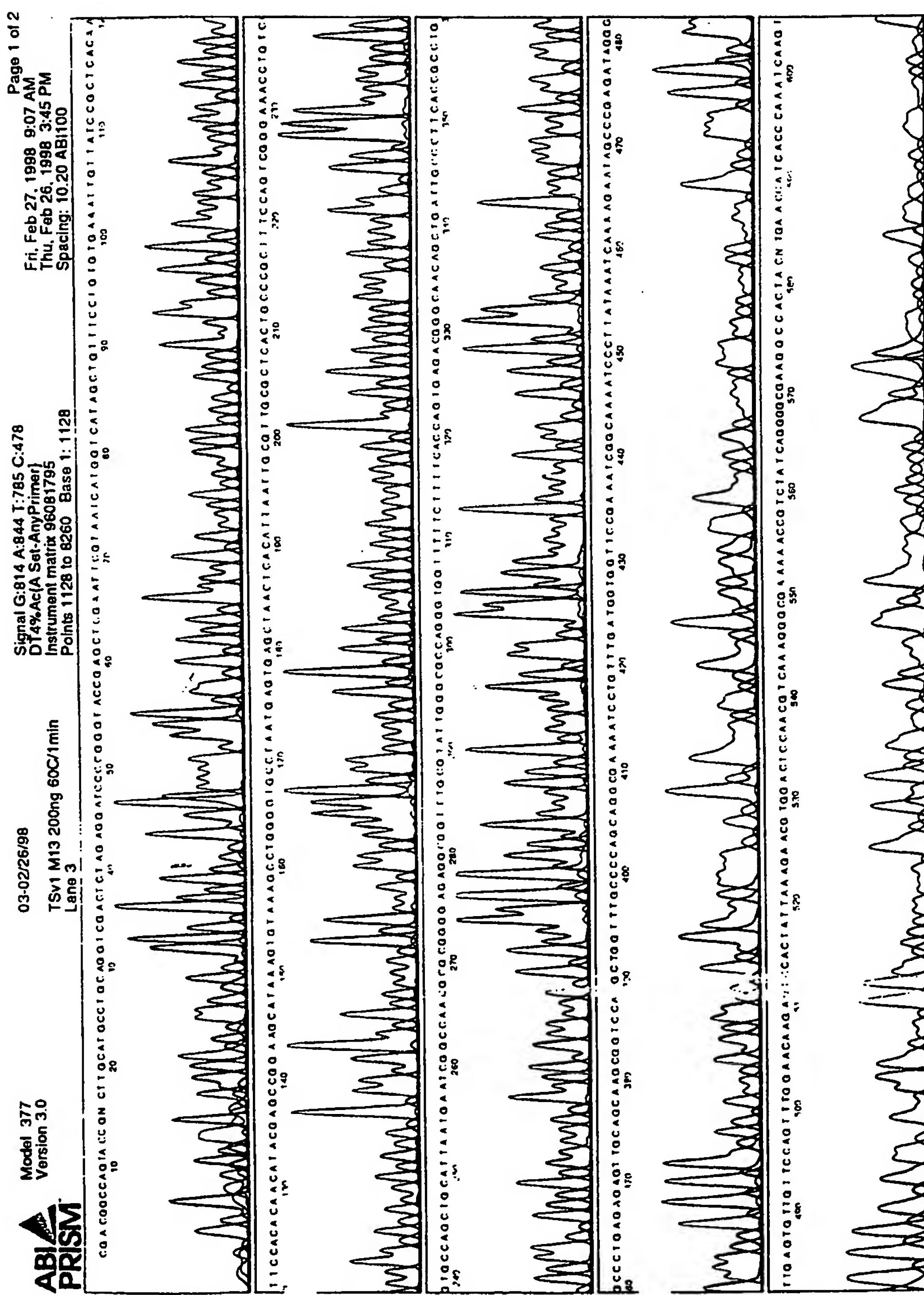


Figure 2c

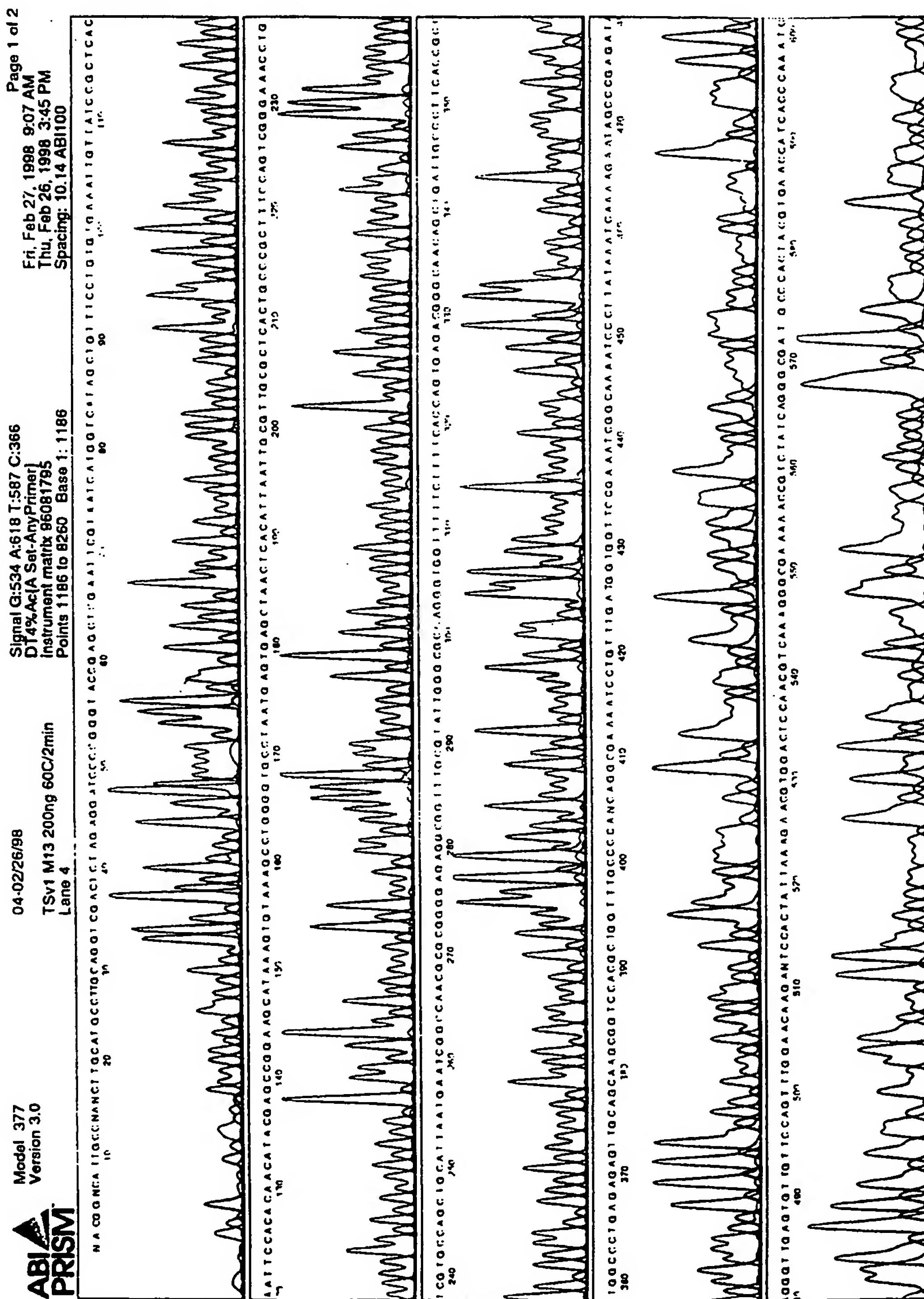
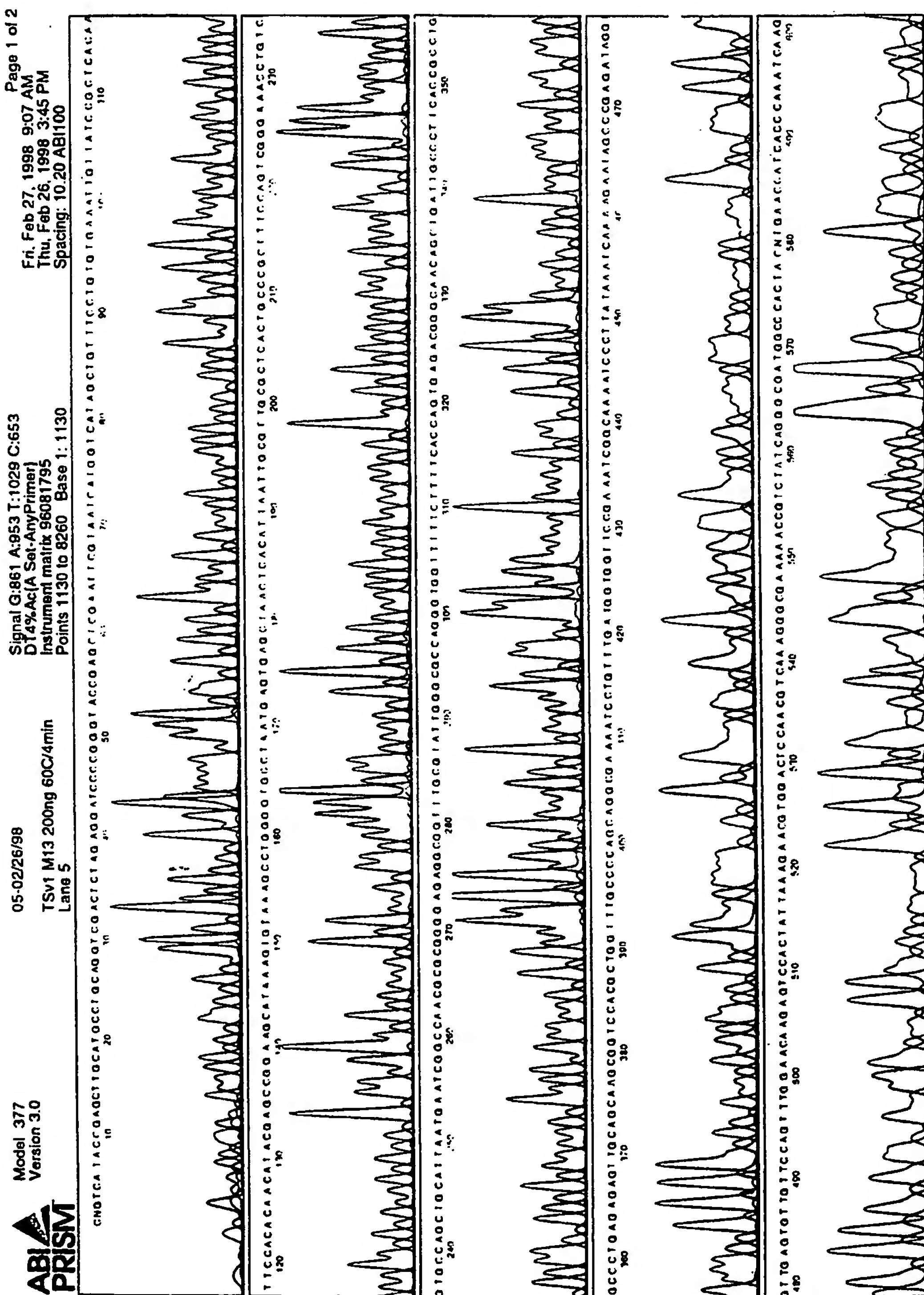


Figure 21



## Figure 2.2

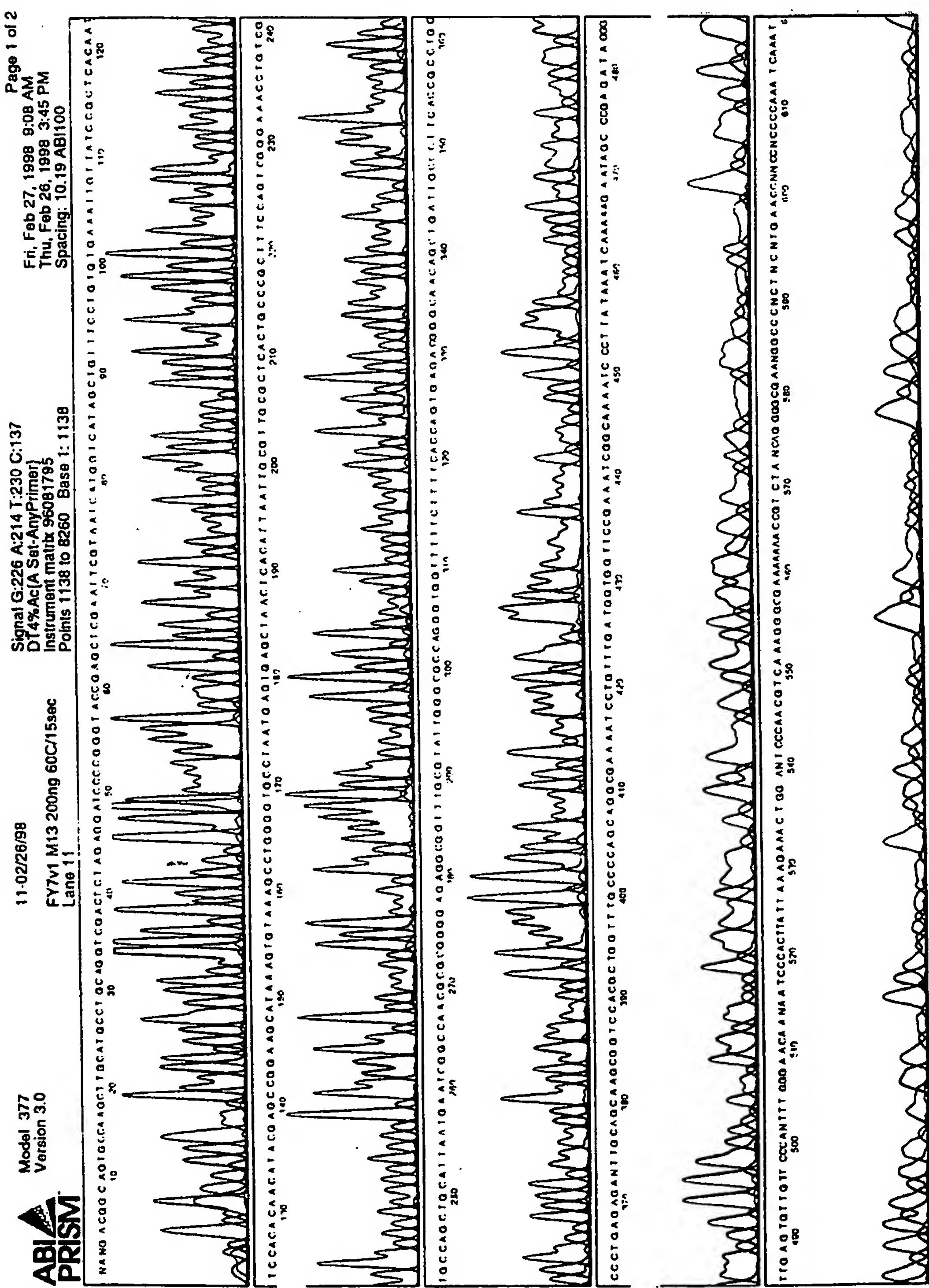


Figure 23

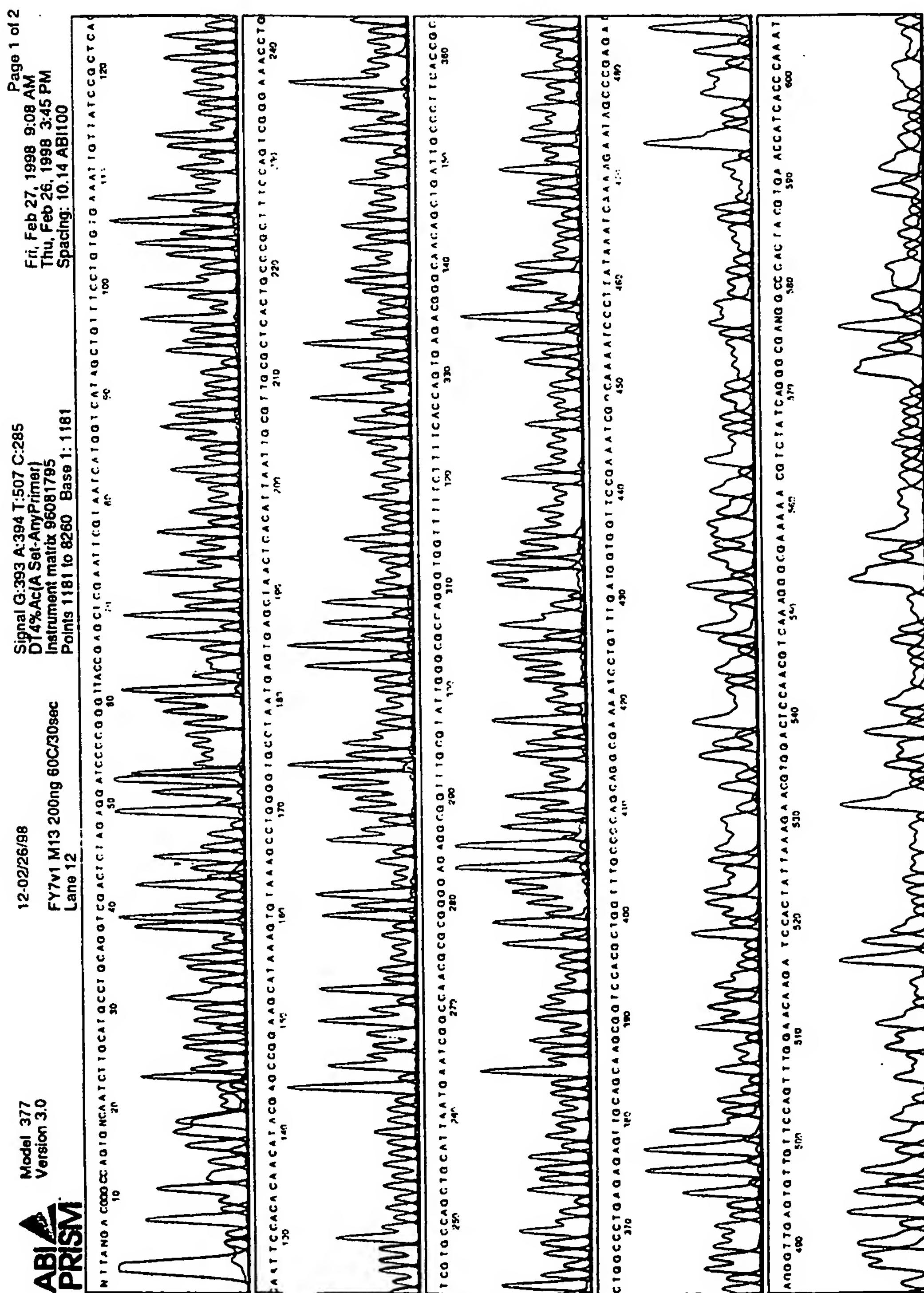


Figure 24

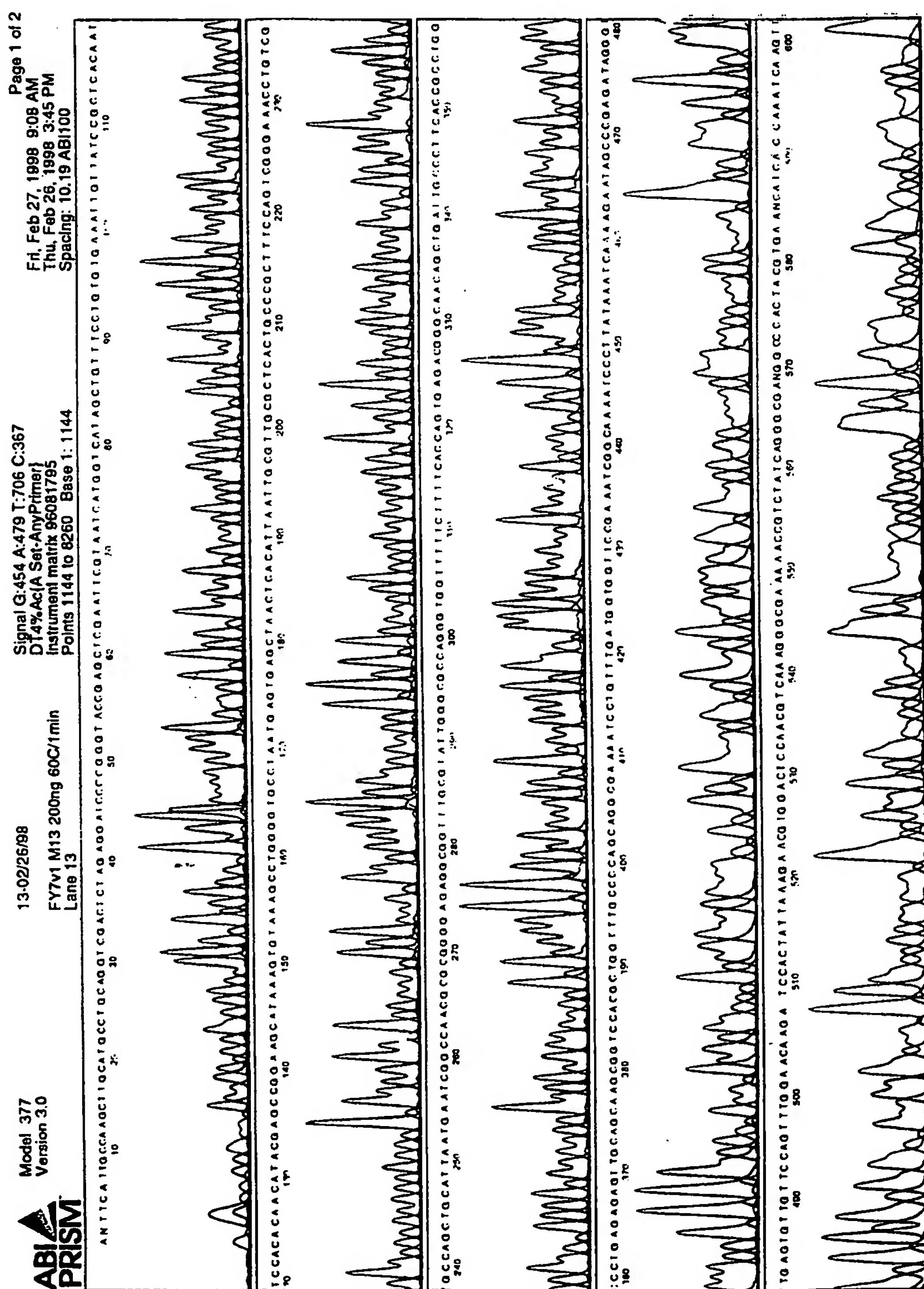


Figure 25

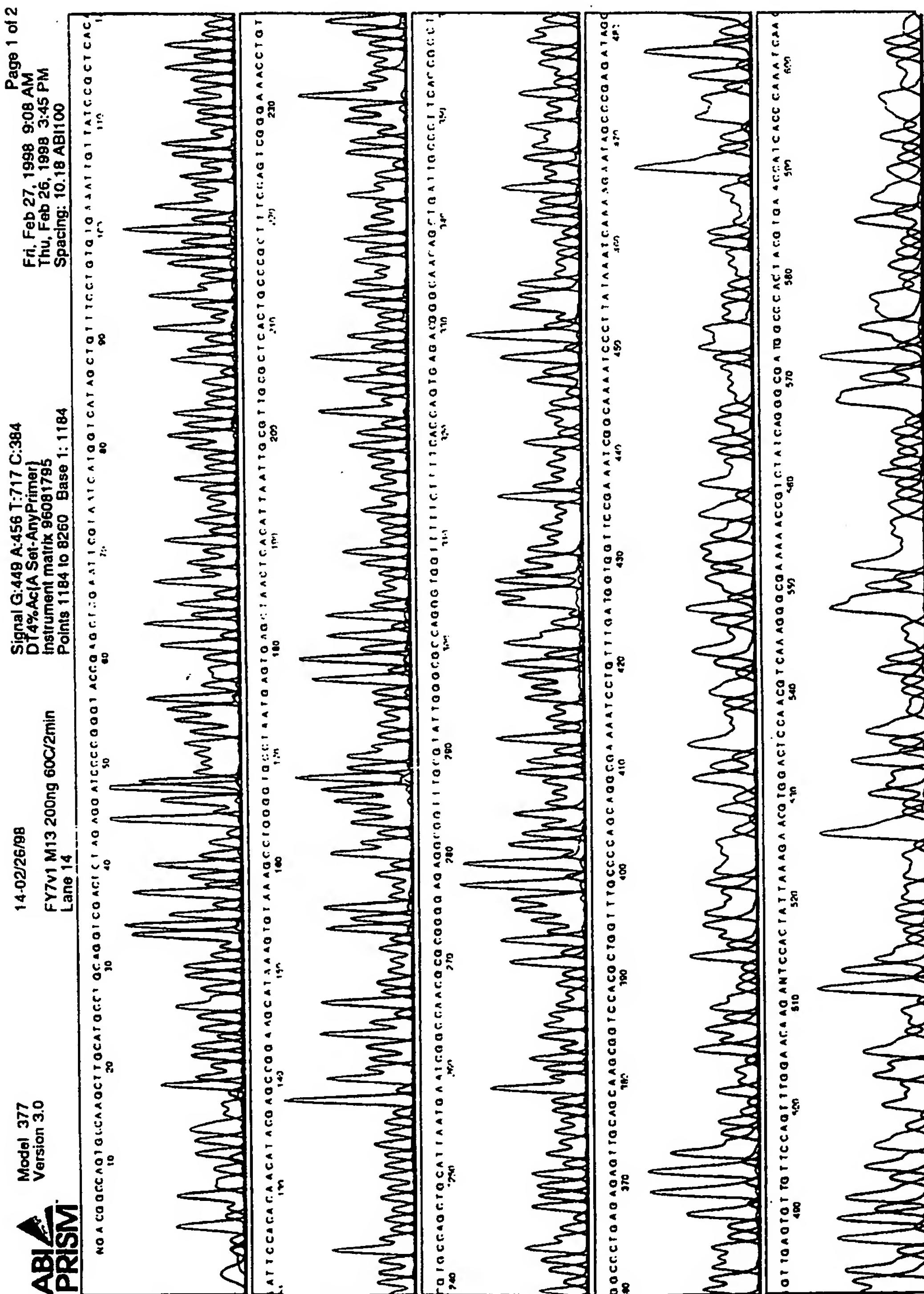


Figure 26

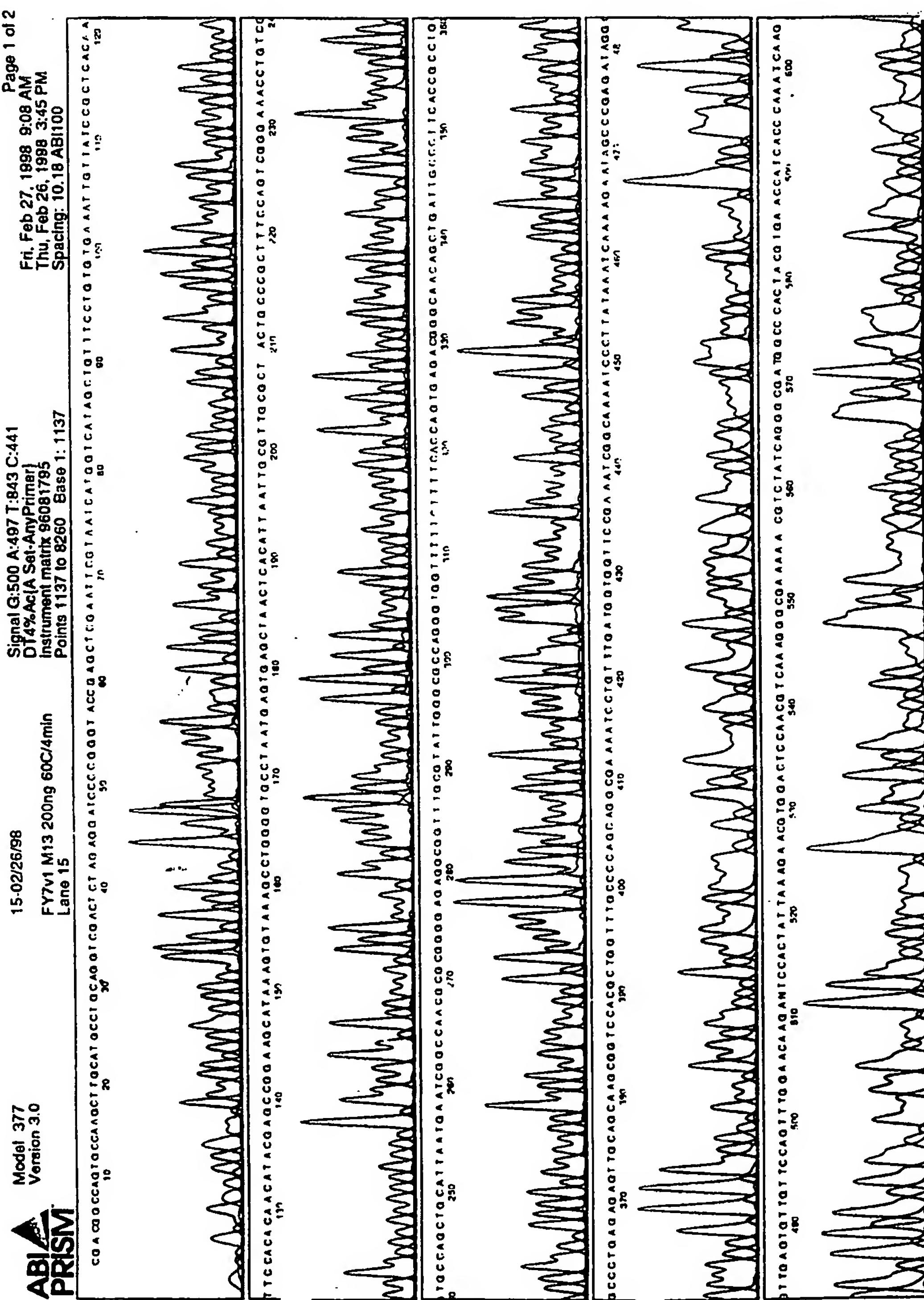


Figure 27